THE ACCOUNTING REVIEW

ucl
@ 12/16/08

VOLUME 83, No. 3 MAY 2008

WILLIAM R. BABER and ANGELA K. GORE
Consequences of GAAP Disclosure Regulation: Evidence from Municipal Debt Issues

KEE-HONG BAE, HONGPING TAN, and MICHAEL WELKER
International GAAP Differences: The Impact on Foreign Analysts

MARY E. BARTH, LESLIE D. HODDER, and STEPHEN R. STUBBEN
Fair Value Accounting for Liabilities and Own Credit Risk

MESSOD DANIEL BENEISH, MARY BROOKE BILLINGS, and LESLIE D. HODDER
Internal Control Weaknesses and Information Uncertainty

JENNIFER L. BROWN and LINDA K. KRULL
Stock Options, R&D, and the R&D Tax Credit

EDDY CARDINAELS and EVA LABRO
On the Determinants of Measurement Error in Time-Driven Costing

DANIEL A. COHEN, AIYESHA DEY, and THOMAS Z. LYS
Real and Accrual-Based Earnings Management in the Pre- and Post-Sarbanes-Oxley Periods

PIERRE JINGHONG LIANG, MADHAV V. RAJAN, and KOROK RAY
Optimal Team Size and Monitoring in Organizations

MARK T. SOLIMAN
The Use of DuPont Analysis by Market Participants

A JOURNAL OF THE
AMERICAN ACCOUNTING ASSOCIATION

AMERICAN ACCOUNTING ASSOCIATION

The members of the 2007–2008 Executive Committee of the Association are as follows:

President	Gary John Previts, Case Western Reserve University
President-Elect	Susan F. Haka, Michigan State University
Vice President	P. Jane Saly, University of St. Thomas
Vice President–Education	Philip M. J. Reckers, Arizona State University
Vice President–Finance	Joseph V. Carcello, The University of Tennessee
Vice President–International	Sidney J. Gray, University of Sydney
Vice President–Professional Relations	Bob Daugherty, PricewaterhouseCoopers
Vice President–Publications	David C. Burgstahler, University of Washington
Vice President–Research	Arnold M. Wright, Northeastern University
Vice President–Sections and Regions	Mark M. Higgins, University of Rhode Island
Past President	Shyam Sunder, Yale University

The American Accounting Association was founded in 1916 as the American Association of University Instructors in Accounting. Its present name was adopted in 1936. The Association is a voluntary organization of persons interested in accounting education, research, and practice. Full membership entitles one to receive Accounting Education News, a selection of one, two, or all three of the Association-wide journals (*The Accounting Review, Accounting Horizons*, and *Issues in Accounting Education*) and the opportunity to purchase Association publications at a discount.

Membership in the Association is available at the following annual rates:

Full membership:
The membership dues include online selection of one, two, or all three of the following Association-wide journals: *The Accounting Review, Accounting Horizons,* and *Issues in Accounting Education.* (This does not include Section journals.)

Select one Association-wide journal	U.S. $165.00	with Print Option $185.00
Select two Association-wide journals	U.S. $175.00	with Print Option $195.00
All three Association-wide journals	U.S. $185.00	with Print Option $205.00

Student membership:
The print option is not available for student members.

Select one Association-wide journal	U.S. $25.00
Select two Association-wide journals	U.S. $35.00
All three Association-wide journals	U.S. $45.00

In addition, full members may belong to one or more Sections: Accounting, Behavior and Organizations; Accounting Programs Leadership Group; American Taxation Association; Artificial Intelligence/Emerging Technologies; Auditing; Diversity; Financial Accounting and Reporting; Gender Issues and Worklife Balance; Government and Nonprofit; Information Systems; International Accounting; Management Accounting; Public Interest; Teaching and Curriculum; and Two-Year College. Each Section assesses its own annual dues, which are collected by the Association.

Editorial correspondence and all manuscripts should be emailed to TARUAZ@email.arizona.edu or sent to Professor Dan Dhaliwal, Senior Editor. *The Accounting Review*, The University of Arizona, Department of Accounting, McClelland Hall 301, 1130 E. Helen Street, Tucson, AZ 85721-0108.

The submission fee of $125.00 for members of the AAA or $200.00 for nonmembers should be paid online (VISA or MasterCard only) at: https://aaahq.org/AAAforms/journals/tarsubmit.cfm and must be paid when article is submitted. Request for reprint permission, subscriptions, advertisements, and dues should be sent to American Accounting Association, 5717 Bessie Drive, Sarasota, FL 34233-2399 • Telephone: (941) 921-7747 • Fax: (941) 923-4093 • Email: Office@aaahq.org. Notice of change of address should be mailed at least eight weeks in advance to ensure undelayed receipt of *The Accounting Review.*

The Accounting Review (ISSN 0001-4826 print and ISSN 1558-7967 online) is published six times a year in January, March, May, July, September, and November by the American Accounting Association, 5717 Bessie Drive, Sarasota, Florida 34233-2399. General subscriptions are available at the annual rates of $340 for a print copy or at $380 for a print and electronic copy. The member subscription rate for a hard-copy is $30. Electronic access is included with membership dues. Periodical postage paid at Sarasota, Florida and additional mailing offices. **POSTMASTER:** Send address change to *The Accounting Review*, 5717 Bessie Drive, Sarasota, Florida 34233-2399.

© 2008 American Accounting Association. All rights reserved. Printed in the United States of America.

THE ACCOUNTING REVIEW

A JOURNAL OF THE

AMERICAN ACCOUNTING ASSOCIATION

Senior Editor
Dan S. Dhaliwal
Editors
John E. Core
Steven J. Huddart
William Kinney
Marlys Gascho Lipe
William E. McCarthy
Mark W. Nelson
Madhav Rajan

45

VOL. 83 MAY 2008 No. 3

WILLIAM R. BABER and ANGELA K. GORE
Consequences of GAAP Disclosure Regulation: Evidence from Municipal Debt Issues 565
KEE-HONG BAE, HONGPING TAN, and MICHAEL WELKER
International GAAP Differences: The Impact on Foreign Analysts 593
MARY E. BARTH, LESLIE D. HODDER, and STEPHEN R. STUBBEN
Fair Value Accounting for Liabilities and Own Credit Risk 629
MESSOD DANIEL BENEISH, MARY BROOKE BILLINGS, and LESLIE D. HODDER
Internal Control Weaknesses and Information Uncertainty 665
JENNIFER L. BROWN and LINDA K. KRULL
Stock Options, R&D, and the R&D Tax Credit 705
EDDY CARDINAELS and EVA LABRO
On the Determinants of Measurement Error in Time-Driven Costing 735
DANIEL A. COHEN, AIYESHA DEY, and THOMAS Z. LYS
Real and Accrual-Based Earnings Management in the Pre- and Post-Sarbanes-Oxley Periods 757
PIERRE JINGHONG LIANG, MADHAV V. RAJAN, and KOROK RAY
Optimal Team Size and Monitoring in Organizations 789
MARK T. SOLIMAN
The Use of DuPont Analysis by Market Participants 823

BOOK REVIEWS, Stephen A. Zeff, Editor

Rick Antle, Frøystein Gjesdal, and Pierre Jinghong Liang (editors), ***Essays in Accounting Theory in Honour of Joel S. Demski*** .... ALFRED WAGENHOFER 855
Alnoor Bhimani (editor), ***Contemporary Issues in Management Accounting*** .... KENNETH A. MERCHANT 856
Jayne M. Godfrey and Keryn Chalmers (editors), ***Globalisation of Accounting Standards*** .... BEGOÑA GINER 859
Nicolas Véron, ***The Global Accounting Experiment*** .... GÜNTHER GEBHARDT 861
Capsule Commentary .... STEPHEN A. ZEFF 863
Editorial Data 864
Editorial Policy and Style Information 865

P31443

Editorial Staff of *The Accounting Review* 2007–2008

SENIOR EDITOR
Dan S. Dhaliwal, The University of Arizona and The University of Auckland
EDITORS
John E. Core
Steven J. Huddart
William Kinney
Marlys Gascho Lipe
William E. McCarthy
Mark W. Nelson
Madhav Rajan
EDITORIAL ASSISTANT
Gail Emily Fey, The University of Arizona

American Accounting Association

EXECUTIVE DIRECTOR
Tracey E. Sutherland
COMMUNICATIONS DEPARTMENT
Beverly Harrelson
Diane Hazard
Lisa Habblitz
Kathy Casper
Judy Cothern
Deirdre Harris
Jan Kovarik (subcontractor)
Lorene Porte (subcontractor)
Anne Dempsey (subcontractor)
OTHER AAA STAFF
Debora Gardner
Peggy Turczyn
Pat Stein
Arlene Oliver
Man Fong
Roland LaTulip
James Szelminski
Diane Leger
Barbara Brady
Kathleen Gwinn
Kathleen McNeil

Editorial Advisory and Review Board

Shannon W. Anderson, Rice University
Anil Arya, The Ohio State University
Ramji Balakrishnan, The University of Iowa
Tim Baldenius, Columbia University
Orie E. Barron, The Pennsylvania State University
Sudipta Basu, Emory University
Anne L. Beatty, The Ohio State University
Daniel A. Bens, The University of Arizona
Jennifer Blouin, University of Pennsylvania
Robert J. Bloomfield, Cornell University
James R. Boatsman, Arizona State University
Christine A. Botosan, University of Utah
Mark T. Bradshaw, Harvard University
Brian J. Bushee, University of Pennsylvania
Jeffrey L. Callen, University of Toronto
William M. Cready, The University of Texas at Dallas
Patricia M. Dechow, University of Michigan and University of California, Berkeley
Mark L. DeFond, University of Southern California
Hemang Desai, Southern Methodist University
Ilia D. Dichev, University of Michigan
Sunil Dutta, University of California, Berkeley
Merle M. Erickson, The University of Chicago
Leslie G. Eldenburg, The University of Arizona
Paul E. Fischer, The Pennsylvania State University
Jennifer Francis, Duke University
Jere R. Francis, University of Missouri–Columbia
Jennifer J. Gaver, University of Georgia
Frank B. Gigler, University of Minnesota, Twin Cities Campus
Dan Givoly, The Pennsylvania State University
Cristi A. Gleason, The University of Iowa
Jonathan C. Glover, Carnegie Mellon University
Michelle Hanlon, University of Michigan
R. Lynn Hannan, Georgia State University
Rachel M. Hayes, University of Utah
D. Eric Hirst, The University of Texas at Austin
Frank D. Hodge, University of Washington
S. Paul Hribar, Cornell University
Christopher D. Ittner, University of Pennsylvania
Bruce W. Johnson, The University of Iowa
Karla M. Johnstone, University of Wisconsin–Madison
Kathryn Kadous, Emory University
Sanjay Kallapur, Purdue University and the Indian School of Business
Steven J. Kachelmeier, The University of Texas at Austin
Bin Ke, The Pennsylvania State University
Elizabeth K. Keating, Harvard University
Inder K. Khurana, University of Missouri–Columbia
Ronald R. King, Washington University in St. Louis
April Klein, New York University
Lisa L. Koonce, The University of Texas at Austin
Ranjani Krishnan, Michigan State University
Linda K. Krull, The University of Texas at Austin
Christian Leuz, University of Pennsylvania
Carolyn B. Levine, Carnegie Mellon University
Robert Libby, Cornell University
Robert C. Lipe, University of Oklahoma
Oliver Zhen Li, University of Notre Dame
Laureen A. Maines, Indiana University
Carol A. Marquardt, New York University
Dawn A. Matsumoto, University of Washington
Edward L. Maydew, The University of North Carolina at Chapel Hill
Richard Mendenhall, University of Notre Dame
Michael B. Mikhail, Duke University
Gregory S. Miller, Harvard University
Lillian F. Mills, The University of Texas at Austin
Partha Mohanram, Columbia University
Donald Moser, University of Pittsburgh
James N. Myers, Texas A&M University
Venky Nagar, University of Michigan
Karen K. Nelson, Rice University
Kaye J. Newberry, University of Houston
Doron Nissim, Columbia University

Patricia C. O'Brien, University of Waterloo
Per M. Olsson, Duke University
Mark E. Peecher, University of Illinois at Urbana–Champaign
Kathy R. Petroni, Michigan State University
Morton Pincus, University of California, Irvine
Joseph D. Piotroski, The University of Chicago
Marlene A. Plumlee, University of Utah
Gordon S. Potter, Cornell University
Grace Pownall, Emory University
Suresh Radhakrishnan, The University of Texas at Dallas
Kannan Raghunandan, Florida International University
Shivaram Rajgopal, University of Washington
Scott A. Richardson, University of Pennsylvania
Darren T. Roulstone, The University of Chicago
Rachel Schwartz, University of Illinois at Urbana–Champaign
K. (Shiva) Sivaramakrishnan, University of Houston
Hollis Skaife, University of Wisconsin–Madison
Richard G. Sloan, University of Michigan
Geoffrey B. Sprinkle, Indiana University
Phillip C. Stocken, Dartmouth College
K. R. Subramanyam, University of Southern California
Hun-Tong Tan, Nanyang Technology University/Singapore
Robert H. Trezevant, University of Southern California
Mark Trombley, The University of Arizona
Mohan Venkatachalam, Duke University
Joseph P. Weber, Massachusetts Institute of Technology
Richard H. Willis, Tulane University
Peter D. Wysocki, Massachusetts Institute of Technology
Hong Xie, University of Illinois at Urbana–Champaign
Xiao-Jun Zhang, University of California, Berkeley

PAST EDITORS

1926–1929	William A. Paton
1929–1943	Eric L. Kohler
1944–1947	A. C. Littleton
1948–1949	Robert L. Dixon
1950–1959	Frank P. Smith
1959–1962	Robert K. Mautz
1962–1965	Lawrence L. Vance
1965–1967	Wendell Trumbull
1968–1970	Charles H. Griffin
1971–1972	Eldon S. Hendriksen
1973–1975	Thomas F. Keller
1976–1978	Don T. DeCoster
1978–1983	Stephen A. Zeff
1983–1987	Gary L. Sundem
1987–1990	William Kinney
1990–1994	A. Rashad Abdel-khalik
1994–1997	Robert P. Magee
1997–2000	Gerald L. Salamon
2000–2002	Linda Smith Bamber
2002–2005	Terry Shevlin

THE ACCOUNTING REVIEW
Vol. 83, No. 3
2008
pp. 565–591

Consequences of GAAP Disclosure Regulation: Evidence from Municipal Debt Issues

William R. Baber
Georgetown University

Angela K. Gore
The George Washington University

ABSTRACT: We compare characteristics of municipal debt issues in states that mandate GAAP for municipalities with issues in states that impose no annual financial disclosure requirements. Cross-sectional comparisons indicate that the use of public (versus private) debt is greater, and municipal debt costs are 14 to 25 basis points lower, in states where GAAP is mandated. Moreover, municipalities in states that impose the GAAP requirement realize lower debt costs following the effective date of the regulation. These results suggest that GAAP requirements reduce municipal borrowing costs. More generally, the evidence indicates that financial reporting regulation reduces contracting costs between borrowers and lenders.

Keywords: *disclosure regulation; municipal debt.*

JEL Classification: *G38; H79; K22; M41; M45.*

I. INTRODUCTION

The Governmental Accounting Standards Board (GASB) was established by the Financial Accounting Foundation in 1984 in response to demand for standard accounting and reporting procedures specifically applicable to state and local governments. The authority of the GASB is well established since 1986 when the AICPA recognized GASB as the primary body for setting generally accepted accounting practices (GAAP) for municipalities. By 2000, 15 U.S. states required their municipalities to file annual financial statements using GAAP as established by the GASB.

The objective of this study is to consider whether state requirements that municipalities use GAAP influence municipal debt financing decisions. Addressing this objective is relevant to the extent that "those who finance government or who participate in the financing

---

The authors are grateful to Arthur Allen, Alex Butler, Ro Gutierrez, Lisa Hersrud, Chris Jones, Lihong Liang, Steve Matsunaga, Wayne Mikkelson, Dale Morse, Terry O'Keefe, Kevin Rich, Mike Rozeff, Chuck Trzcinka, Michael Williams, and Jean Zhang for their comments and advice. Workshop participants at the Financial Accounting Foundation, George Mason University, Indiana University–Indianapolis, the University of Oregon, and the 2005 AAA Government Non-Profit section conference on nonprofit financial stewardship at the Kennedy School of Government, Harvard University offered useful suggestions.

Editor's note: This paper was accepted by Dan Dhaliwal.

Submitted March 2006
Accepted October 2007

process" are among the users of state and local government financial reports that GASB is obligated to serve (GASB 2005). In addition, because compliance with GAAP is costly, if tangible benefits of GAAP regulation do not accrue, then decisions to impose GAAP reporting apparently contradict the public interest.[1]

Implications of the study extend beyond the municipal sector. Specifically, the current status of municipal reporting facilitates investigation of the consequences of mandatory disclosure requirements in a relatively controlled environment, as not all states regulate municipal disclosure. Thus, the investigation potentially informs deliberations of more general issues about the consequences of regulating accounting and disclosure practices.

The empirical analysis compares municipal debt issues in states that require GAAP (designated GAAP states) with issues in states that do not regulate municipal disclosure (non-GAAP or unregulated states). The underlying proposition is that, if state GAAP requirements are associated with lower debt financing costs, then the characteristics of debt issues should differ between GAAP states and non-GAAP states.

We begin by investigating whether the relative use of municipal debt differs between GAAP and non-GAAP states. If GAAP disclosure regulation reduces the costs of contracting between municipal borrowers and lenders, then we expect greater use of debt by municipalities in GAAP than in non-GAAP states. Notwithstanding this expectation, we find relatively little evidence of differences between GAAP and non-GAAP states in the use of debt financing.

Next we consider whether municipalities issue debt publicly or negotiate arrangements with private lenders (e.g., banks, insurance firms, or pension funds). We expect that GAAP has greater value in public debt markets, where many borrowers and lenders compete, than in privately negotiated arrangements. This expectation is guided by studies of the use of debt in the private sector context, which predict and find greater proportions of privately negotiated lending arrangements when disclosure costs are high (Dhaliwal et al. 2005). Thus, we anticipate greater use of public versus private debt (e.g., bank debt or private placements) financing by municipalities in GAAP states than in non-GAAP states. The empirical analysis indicates greater use of public debt by municipalities in GAAP states, which suggests that reductions in municipal debt costs associated with state mandated GAAP are greater for public than for private debt issues.

We then compare debt costs (bond yields) for a sample of 424 municipal debt issues, 212 from the GAAP states and 212 from the non-GAAP states. After controlling for market- and issue-specific factors that affect municipal yields, including voluntary disclosure, we find that mean municipal debt costs are lower by 14–25 basis points in states that mandate GAAP.

Finally, for two states that recently imposed GAAP regulation, Montana and Nevada, we compare municipal debt costs before and after GAAP is mandated. Although the sample of 33 debt issues is relatively small, we find that municipal debt costs are significantly lower following the GAAP requirement. This result corroborates the cross-sectional comparisons of municipal debt costs in GAAP states versus non-GAAP states.

Three prior studies examine whether state regulation of municipal disclosure affects municipal bond interest costs. Ingram and Copeland (1982) document negative associations between ordinal measures of state requirements and the change in offering yield premiums

---

[1] The analysis is confined to measurable financial costs and benefits to the municipality, and as such, ignores social welfare issues. The focus on costs and benefits in this context is consistent with both GASB-commissioned focus groups and characterizations in prior research that bond market participants are primary users of GAAP financial statements (Gore 2004).

(i.e., secondary market *Blue List* quotes less the corresponding U.S. Treasury yields). In an extension, Benson et al. (1984) find no statistically significant associations between bond yields and state regulation of accounting for a sample of new issues. The authors do find, however, that associations between bond yields and accounting-based measures of financial viability are lower for municipalities in states that impose accounting or financial reporting requirements.[2]

In contrast with Ingram and Copeland (1982) and Benson et al. (1984), Fairchild and Koch (1998) find that yields on primary market municipal bond issues are lower in states where so-called *Blue Sky Laws* require "official statement" filings in anticipation of municipal debt issues. Official statements are comparable to prospectus filings required by the SEC for public security issues in the private sector. Analysis reported in the study indicates that net interest costs are 14 basis points lower for municipalities in states that require official statement filings, compared to those in states that do not require such filings.

The current study contributes to prior literature that examines whether state disclosure regulation affects municipal debt costs in at least four ways. First, we go beyond documenting associations between state disclosure requirements and municipal interest costs to investigate whether disclosure regulation influences both the issuance of municipal debt and the use of public debt markets versus private placement. These associations have not been investigated previously.

Second, we find that state-mandated GAAP requirements contribute incrementally toward determining municipal debt characteristics beyond other state regulations designed to monitor or control municipal financing decisions, including official statement regulations considered by Fairchild and Koch (1998). In contrast, prior studies do not consider the effects of alternative state regulations or state characteristics that can also affect municipal debt costs.

Third, the study advances the literature methodologically. Specifically, we address a larger sample that includes debt characteristics not considered in prior studies, and compare mean municipal debt costs surrounding the implementation of GAAP disclosure regulation. These analyses more comprehensively and definitively establish a role of GAAP accounting requirements in the context of municipal debt markets than prior studies that employ smaller samples and consider cross-sectional associations between debt costs and reporting requirements.

Finally, we use more recent data than prior studies that primarily investigate municipal debt costs during the 1970s and 1980s. New federal regulations that encourage voluntary disclosure—for example, the Single Audit Act—potentially obviate the need for state disclosure regulation. On the other hand, the involvement of the GASB has led to a richer and more comprehensive set of GAAP requirements, which can strengthen the relations found in prior studies (Reck et al. 2004; Hildreth and Zorn 2005).[3] Thus, the current municipal reporting environment differs substantially from periods investigated in prior studies and, therefore, the relation between GAAP disclosure regulation and municipal bond interest costs using recent data is not obvious.

---

[2] In another study, Benson et al. (1991; hereafter, BMR) find that, for states with substantial short-term debt, general obligation bond interest costs are lower for states that disclose GAAP information. GAAP disclosure by state governments is voluntary rather than imposed. In contrast, we focus on consequences of state regulation of municipal disclosure. Our study also differs from BMR in its focus on municipalities, rather than states. BMR (1991, 309) note that empirical results for state government issues often differ from results for municipal issues.

[3] As examples, newly implemented disclosure standards—including standards regarding investment disclosure (GASB Nos. 3 and 31), pension disclosure (GASB Nos. 5, 25, and 27), risk financing disclosure (GASB No. 10), and a standard financial reporting model (GASB No. 34)—are potentially useful for assessing risk.

As a whole, the analysis supports the proposition that state-imposed requirements on municipalities to use GAAP accounting are associated with lower municipal debt costs. Moreover, the evidence suggests that the consequences of state-imposed GAAP tend to be more substantial for small municipalities. This evidence should be interpreted cautiously, however, as we do not consider the costs of disclosure regulation, which potentially exceed benefits realized in debt markets. For example, while we find that differences in debt costs in GAAP states tend to be more substantial for smaller municipalities, disclosure costs are likely to be more substantial as well.

Section II provides institutional background and poses three hypotheses regarding the relation between disclosure regulation and municipal debt characteristics. Sample selection and measures used in the analysis are in Section III. Results for multivariate specifications are in Section IV. Concluding remarks are in Section V.

II. HYPOTHESIS DEVELOPMENT

Background

Prior to the 1970s, oversight of municipal accounting and bond issuance was inconsistent and frequently ineffective (Gellis 1996). High-profile defaults by New York City in the 1970s, the Washington Public Power Supply System (WPPSS) in the 1980s, and Orange County in the 1990s brought municipal accounting under scrutiny and created a demand for regulation to remedy alleged abuses. In response, financial disclosure and audit requirements were implemented at both the federal and state levels.

Consistent with state sovereignty guarantees in the U.S. Constitution, the 1933 and 1934 Securities Acts specifically exempt municipal securities from federal regulation. The U.S. Congress reaffirmed this proscription regarding federal government oversight of municipal debt issues in 1975 by passing the Tower Amendment, which specifically prohibits the Securities and Exchange Commission (SEC) from requiring issuers of municipal securities to file with the SEC (Hildreth and Zorn 2005). The trading environment for U.S. municipal debt, which exceeded $2 trillion at the beginning of 2005 (Hildreth and Zorn 2005), therefore differs in important ways from that for corporate debt. More specifically, unlike publicly traded securities issued by private sector firms, securities issued by state and local governments—and more to the point of this study, accounting and reporting practices used by municipalities that issue publicly traded debt—are not directly regulated by the federal government.

The federal government can influence municipal financial reporting practices indirectly, however, and has increasingly done so since New York City's 1975 financial crisis (Soybel 1992; Reck and Wilson 2006). A salient example is the Single Audit Act of 1984, which requires audited financial statements for municipalities receiving federal funding in excess of $500,000.[4] While the Act requires audits for certain large municipalities, it does not require specific municipal accounting and disclosure methods, including GAAP standards, which remain the purview of state regulators. In fact, non-GAAP states can require municipalities in the state to follow alternative accounting methods, including the cash basis.[5]

---

[4] The original Act applies to municipalities receiving federal funding greater than $100,000. The threshold was increased to $300,000 in 1996 and to $500,000 in 2003 (Reck and Wilson 2006).

[5] To illustrate, consider the following excerpt from the U.S. Department of Housing and Urban Development (HUD): "Neither the Act nor A-128 require financial statements to be presented in accordance with Generally Accepted Accounting Principles (GAAP). If a government entity reports on the cash basis, those reports, with an appropriate auditor's opinion, will satisfy the single audit requirements" (HUD 1984).

Moreover, the SEC issued rule 15c2-12 in 1994, exploiting anti-fraud provisions in the 1934 Securities Act. The Rule requires brokers, dealers, and underwriters to obtain commitments from municipal borrowers in excess of $1 million to disclose financial information annually (Fairchild and Koch 1998; Reck and Wilson 2006). Rule 15c2-12 is clearly intended to force municipalities to disclose information to secondary bond market participants. The underlying legislation does not provide for an enforcement mechanism, however, and as a result, many municipalities do not comply (Hume 2002), and brokers, dealers, and underwriters lack incentives to force compliance (Gellis 1996).

Although neither of these federal regulations requires municipalities to use GAAP financial reporting practices, they do potentially encourage voluntary GAAP compliance by municipalities in states where disclosure is unregulated. These recent changes in the municipal financial reporting environment raise questions about whether state-imposed GAAP disclosure regulations are necessary or effective.

In contrast with the federal government, states can directly dictate financial reporting practices for municipalities (Benson et al. 1991). Some states apparently adopted GAAP following the high-profile bankruptcy filings of New York City of $600 million in 1975 and the WPPSS of $2.5 billion in 1988, as in both of these cases, inadequate financial disclosure was cited as a contributing factor (Gellis 1996). Coincidentally or not, several states adopted GAAP regulations within five years following the New York City default (i.e., Rhode Island, Florida, and Kentucky), while two others (Connecticut and Ohio) adopted GAAP shortly after the WPPSS default. On the other hand, some states (e.g., Massachusetts and New Jersey) impose state-specific accounting practices that differ from methods established and promulgated by the GASB, while other states require GAAP for some, but not all, municipalities.[6] Finally, some states—for example, Texas—impose no disclosure requirements on municipalities.

Addressing why some states require GAAP while other states require virtually nothing is beyond the scope of this study; however, some non-GAAP states appear reluctant to mandate disclosure due to cost considerations. For example, communication with the authors suggests that the Iowa does not require GAAP because they "have a large number of small communities [that] would not see any benefit from this requirement." Another state regulator suggests that municipalities that issue bonds have incentives to voluntarily comply with GAAP, while smaller municipalities, which seldom issue debt, could avoid the cost of GAAP compliance. Such reasoning is used to justify piecemeal GAAP requirements in some states where population, revenue, or debt thresholds determine which municipalities are exempted from GAAP requirements.

Regardless of specific state requirements, and in contrast with federal regulators, state regulators can and do enforce compliance with state-imposed accounting and disclosure regulations. For example, in Michigan, which requires GAAP, state auditors routinely test GAAP compliance by reviewing municipal reports. If a financial report is not in compliance, then the municipality must submit a revised financial report (Gore 2004).

To summarize, in contrast with markets for most publicly traded securities where federal regulation dominates, state governments primarily establish and enforce municipal accounting and disclosure practices.[7] Thus, information transparency and disclosure in municipal debt markets varies substantially and is typically less than that for private sector debt issues.

---

[6] New Jersey, which accepts reports from municipalities using three different bases of municipal accounting, recently debated whether to begin requiring GAAP for all municipalities, citing comparison within the state and with municipalities in other states as a motive for doing so (Higginbotham 2001).

Consequences of Disclosure Regulation

Prior studies, set primarily in the context of federally regulated securities markets, suggest that borrowers have private incentives to disclose voluntarily when voluntary disclosure is beneficial to contracting parties (Leftwich 1980; Easterbrook and Fischel 1996; Healy and Palepu 2001). Empirical studies of factors that influence the cost of debt financing in the private sector indicate inverse associations between voluntary disclosure and the cost of capital (Sengupta 1998; Botosan and Plumlee 2004). Such results raise questions about whether regulation is justified when accounting and disclosure practices can be used voluntarily.

On the other hand, Bushee and Leuz (2005) suggest circumstances where even voluntary disclosers benefit from disclosure regulation. First, regulation pre-commits regulators to enforcement and borrowers to future disclosure. Lacking such commitment, voluntary disclosers can withhold information opportunistically. If lenders demand compensation for the risk that borrowers behave opportunistically, then borrowers potentially benefit from disclosure requirements. A second benefit of disclosure regulation arises when information availability influences the overall search and transaction costs, and thus influences liquidity, in lending markets. If disclosure improves market liquidity, then decisions to disclose voluntarily fail to incorporate public benefits that accrue to other lenders from liquid capital markets. That is, voluntary disclosure decisions fail to consider "information spillover," a positive externality that accrues when borrowers act collectively. If either of these scenarios applies, then disclosure regulation potentially provides benefits beyond those that result from voluntary disclosure.

Requiring public financial disclosure has potentially adverse consequences beyond the out-of-pocket bookkeeping costs of complying, as requirements to report information publicly can increase agency costs of an organization in important, but subtle, ways. As examples, the disclosure of proprietary information in the private sector can undermine an organization's position with respect to competitors or contracting with other external parties (Bowen et al. 1995). Moreover, Dhaliwal et al. (2005) document positive associations between the relative use of private (versus public) debt financing and measures that suggest high costs of public disclosure. Private financing typically does not require public financial disclosure; therefore, such evidence supports the premise that public disclosure is costly. Further, managers can respond to disclosure regulation by reducing reliance on other less costly mechanisms that could be implemented voluntarily in an unregulated environment. For example, Gore et al. (2004) describes the use of bond insurance as a substitute for public disclosure to reduce municipal debt costs. Thus, notwithstanding evidence that voluntary disclosure reduces the costs of debt financing, the consequences of mandating financial disclosure are not straightforward.

GAAP Disclosure Regulation and Decisions to Use Municipal Debt Financing

Potentially important differences between the public and private sectors suggest caution when importing results from publicly traded companies into the context of municipal debt. As examples, existing studies posit relatively less liquidity (Downing and Zhang 2004) and higher transaction costs (Harris and Piwowar 2006) in municipal bond markets, which implies greater information asymmetry than in corporate debt markets. Moreover, municipal governments do not compete for market share in the usual sense (although see Tiebout 1956), so municipal managers are less likely than their private sector counterparts to be concerned about disclosure of proprietary information. These differences suggest that net benefits to disclosure regulation are greater in the context of municipal debt markets than in private debt markets.

---

[7] In addition to annual disclosure requirements, states can impose other municipal regulations such as audits or official statement disclosures. We discuss implications of such alternatives later.

On the other hand, public disclosure can adversely influence transactions costs incurred under implicit contracts, such as those with suppliers and employees (as in Bowen et al. 1995). Thus, predicting how disclosure regulation influences the use and characteristics of municipal debt requires thinking comprehensively about how disclosure regulation affects decisions about debt-financed municipal spending. More specifically, a municipality's use of debt financing can be endogenous to the financial reporting environment in the sense that the marginal cost of debt influences incentives to undertake debt-financed capital projects. If GAAP disclosure regulations are associated with lower debt costs, then municipalities can respond by undertaking more activities that use debt financing. Given this scenario, we expect greater use of debt financing in states that impose GAAP.

Alternatively, since GAAP requirements are only one of a variety of potentially less costly mechanisms that states can impose to reduce municipal financing costs, state requirements that municipalities follow GAAP may not affect municipal incentives to use debt financing. This might be the case, for example, if alternative mechanisms reduce adverse selection costs to the efficient minimum such that GAAP reporting provides lenders with no incremental assurance (Gore 2004; Gore et al. 2004). Thus, the effect of state-mandated GAAP on the use of municipal debt is not straightforward. The related null hypothesis is:

H1: The issuance of municipal debt is independent of whether GAAP is required by the state government.

Next we consider that municipal debt financing can be either issued publicly or negotiated with private lenders. Although debt-related savings from GAAP potentially accrue to municipalities that engage all types of lenders, the benefits of GAAP are likely greater in the context of public markets where the value of public debt is determined by many actual and prospective lenders. Municipalities in unregulated states can, at some cost, disclose GAAP information to public lenders voluntarily, but municipalities in states that require GAAP incur no incremental cost of disclosure. Thus, if GAAP disclosure regulation reduces the costs of contracting with public lenders, then we expect relatively greater use of private than public debt financing in non-GAAP states than in GAAP states. This characterization suggests the following null hypothesis:

H2: The relative use of private (versus public) municipal debt financing is independent of whether GAAP is required by the state government.

Rejecting H1 or H2 in favor of finding greater use of municipal debt or relatively greater use of public debt in GAAP states supports the proposition that the GAAP requirement reduces the marginal cost of municipal debt financing. On the other hand, not rejecting these hypotheses, or rejecting in favor of the alternative hypotheses that the use of debt financing or public debt is less for GAAP states, implies that GAAP regulation is not an important determinant of municipal debt costs beyond what is provided through voluntary disclosure or by other state requirements.

GAAP Disclosure Regulation and the Cost of Debt Financing

Notwithstanding evidence from prior literature that voluntary disclosure reduces state government bond interest costs (Benson et al. 1991), associations between GAAP regulation and mean public debt costs are not obvious. One perspective is that the marginal cost of new debt is unchanged when states mandate GAAP because federal regulations encourage adequate disclosure indirectly among municipalities in non-GAAP states. Moreover, if regulation reduces marginal debt costs, then municipalities may pursue additional debt financing to the point where cross-sectional differences between average debt costs for regulated versus voluntary regimes are not detectable.

On the other hand, state-mandated GAAP potentially assures lenders of continuing disclosure and enforcement and/or promotes municipal debt market liquidity such that benefits to municipal lenders and borrowers can go beyond what exists when GAAP disclosure is voluntary. If so, then we anticipate lower municipal debt costs when GAAP is state-mandated. The following null hypothesis applies:

> **H3:** The cost of municipal debt financing is independent of whether GAAP is required by the state government.

Rejecting the null H3 in favor of lower debt costs in GAAP states supports the premise that GAAP disclosure regulation reduces average municipal debt costs.

The GAAP Requirement as Part of a Comprehensive State Governance Policy

Finally, we recognize that GAAP disclosure is only one of many mechanisms that state governments can use to oversee and control the actions of municipalities and assure prospective lenders regarding the risk of debt issues. Thus, the GAAP requirement potentially distinguishes municipalities where the state government centralizes fiscal control over municipalities. If so, then differences between GAAP and non-GAAP states can indicate the consequences of comprehensive state policies that include mandated GAAP, rather than direct effects of the GAAP requirement itself.[8]

To address these possibilities, we consider statutory differences among the states that potentially influence contracting between municipal borrowers and lenders. More specifically, following the international governance literature (La Porta et al. 1998), we consider measures of regulation, creditors' rights, and enforcement. To consider regulation, we identify states that mandate official statement requirements (Fairchild and Koch 1998). To consider creditor rights, we ascertain whether state law proscribes municipal bankruptcy. We use three indicators of enforcement. Specifically, we consider whether municipal audits are mandated, whether the percentage of CPAs on the state auditor's staff is above the median, and whether the state has a corruption index that exceeds the sample median.[9]

---

[8] This point is also relevant in the context of international studies. La Porta et al. (1997) and La Porte et al. (1998) conjecture that "strong" financial disclosure laws potentially compensate for relatively weak investor protection laws. Consistent with this conjecture, Miller and Puthenpurackal (2002) find that yields on public debt offerings in the U.S. by foreign firms (or "Yankee" bonds) are greater by 52 basis points for issues from countries with relatively weak investor protection laws. As in the international context, disclosure requirements vary among states, but in contrast with international markets, the municipal debt market has the same general underlying system of investor protection laws. In this regard, the consequences of accounting regulation are potentially easier to document in the municipal debt context.

III. SAMPLE AND DATA

Following Gore and Trzcinka (2004), who survey each state to determine the municipal disclosure regulations, we identify 15 states that require municipalities to prepare annual GAAP-based financial statements (regardless of whether they issue debt) and ten states that do not require annual financial reporting by municipalities.[10] The conceptual argument that supports a role for state regulation beyond voluntary compliance presumes that the GAAP requirement be irrevocable so that future investors in municipal bonds can rely on GAAP reports into the future. Thus, to be classified as a GAAP state, the GAAP requirement must be imposed by state law. If the GAAP requirement is revocable, then the requirement has no more economic substance than a rule that simply permits municipalities to use GAAP voluntarily. For example, if an executive or legislative oversight group within the state government imposes the requirement, then the next administration or the next finance committee can change the rule. Thus, a rule imposed by a state government administrator does not carry the same weight as a requirement imposed by state statute.[11]

Table 1 shows state size (1998 population), state marginal personal income tax rates, and the number of municipal bond issues for each GAAP state (Panel A) and each unregulated state (Panel B). The remaining 25 states require piecemeal, hybrid, or inconsistent state-specific disclosure regulations. As a result, state disclosure requirements for these 25 states are ambiguous and contain relatively high cross-sectional variation such that comparisons with GAAP states or unregulated states are not meaningful. We therefore restrict the primary investigation to municipalities within the states that clearly and definitively either require or do not require GAAP for all municipalities.[12]

We classify sample observations based on state requirements, not individual municipal disclosure practices. Municipalities can, and frequently do, disclose financial information voluntarily—including information not required by GAAP—and therefore, the extent of *actual* financial disclosure by specific municipalities in either GAAP or unregulated states both varies and potentially exceeds GAAP standards (Gore 2004). Thus, we control for individual municipal disclosure in the multivariate tests.

---

[9] The corruption index indicates indictments of state and local officials. See Butler (2004) for details.

[10] The Gore and Trzcinka (2004) survey supplements a 2000 survey conducted by the National Association of State Auditors, Comptrollers, and Treasurers (NASACT). The NASACT survey is often incomplete and/or apparently contradicts a 1996 NASACT survey. NASACT surveys also do not distinguish unregulated states from states that require state-specific disclosure, nor do they distinguish regulations for municipalities from regulations for other types of government entities (schools, counties, and authorities). Thus, we rely primarily on the Gore and Trzcinka (2004), rather than the NASACT, classification. We validate the Gore and Trzcinka (2004) classification by searching relevant state statutes for evidence of municipal disclosure requirements. One state, Virginia, is thus removed from the sample due to concerns about the classification. Results using alternate classifications, including straightforward use of the NASACT classifications (i.e., 1996, 2000, 2002), do not differ qualitatively from those reported.

[11] One state, Virginia, appears to impose GAAP administratively rather than statutorily. Virginia is unusual in that, although statutes do not require GAAP, the statutes assign the Auditor of Public Accounts (a state government administrator) with the responsibility to impose financial reporting requirements on municipalities. Thus, a case can be made that Virginia is a non-GAAP state. The Auditor of Public Accounts requires GAAP for Virginia municipalities with population greater than 3,500. Thus, we treat Virginia as a piecemeal state and omit it from the analysis reported later. Results are comparable when Virginia is included as the eleventh non-GAAP state.

[12] Municipalities in Montana and Nevada, which imposed the GAAP requirement during 2001, are excluded and analyzed separately.

TABLE 1
Characteristics of States that Require GAAP Reporting (GAAP States) and States that do not Impose Reporting Requirements (Non-GAAP States)[a]

State	Population	Highest Marginal Personal Income Tax Rate	Number of Bond Issues Outstanding
Panel A: GAAP States (n = 15)			
Arizona	4,667,277	5.04	812
Connecticut	3,272,563	4.50	750
Florida	14,908,230	0.00	3,300
Georgia	7,636,522	6.00	1,410
Kentucky	3,934,310	6.00	740
Michigan	9,820,231	4.20	2,392
Montana[b]	879,533	11.00	151
Nevada[b]	1,743,772	0.00	550
New Hampshire	1,185,823	5.00	348
New Mexico	1,733,535	8.20	456
North Carolina	7,545,828	7.75	846
Ohio	11,237,752	6.98	1,332
Rhode Island	987,704	9.90[c]	264
South Dakota	730,789	0.00	148
Tennessee	5,432,679	6.00	1,092
Panel B: Non-GAAP States (n = 10)			
Alabama	4,351,037	5.00	862
Delaware	744,066	5.95	144
Idaho	1,230,923	8.20	294
Missouri	5,437,562	6.00	858
Nebraska	1,660,772	6.68	252
North Dakota	637,808	5.54	176
Pennsylvania	12,002,329	2.80	3,515
South Carolina	3,839,578	7.00	832
Texas	19,712,389	0.00	5,633
Vermont	590,579	9.50[c]	85

---

[a] GAAP states require municipalities to prepare financial statements that comply with Generally Accepted Accounting Practices (GAAP) as specified by the Governmental Accounting Standards Board (GASB). Non-GAAP states impose neither GAAP nor other state-specific disclosure regulations. The remaining states, which are not shown, impose hybrid and/or state-specific disclosure requirements. Population is the 1998 U.S. Census Bureau estimate. Highest Marginal Personal Income Tax Rate is for 1998 per the Federation of Tax Administrators (available at: http://www.taxadmin.org). Number of Bond Issues Outstanding is from the 1998 Moody's bond record.

[b] Montana and Nevada imposed GAAP in 2001, which is during the period encompassed by the analysis. Thus, observations for these states are omitted from cross-sectional tests and analyzed separately.

[c] Rhode Island and Vermont's marginal rates are 25.5 percent and 24 percent of federal taxable income, respectively.

We use data from several sources. Hypothesis 1 considers the amount of total debt issued (both public and private) for each municipality. These data are obtained from the

U.S. Bureau of Census (Census data) for 1995 to 2002, resulting in 4,955 observations.[13] Data for individual municipal disclosure are from the Government Finance Officer's Association (i.e., the GFOA certificate of achievement).

Hypothesis 2 involves the relative use of public versus private debt. These details are from the Thomson Financial SDC Platinum database (SDC data). Combining SDC data with Census data yields 2,049 observations with complete information.

Addressing H3 requires details of public bond issues. We obtain complete information for 1,849 new bond issues (that is, bonds issued in the primary market) in GAAP and unregulated states during 1995 to 2002 from the SDC data. Observations are either general obligation or revenue bonds issued in the primary market by municipal governments (cities, towns, villages, or boroughs). We focus on primary market bond issues because municipalities have direct incentives to reduce the costs of debt at the original issue. In contrast, debt costs in secondary market transactions are borne by market participants, and therefore municipalities lack direct incentives to minimize secondary market yields. To mitigate concern about statistical independence of the sample observations, we select the most recent bond issue for each of 424 municipalities in the GAAP or unregulated states (Sengupta 1998). These procedures yield 212 public debt issues in GAAP states and 212 issues in non-GAAP states.[14]

State-imposed auditing requirements and the percentage of CPAs on the state auditor's staff are from National Association of State Auditors, Comptrollers, and Treasurers (NASACT 1996, 2000, 2002); official statement requirements are from Fairchild and Koch (1998); municipal bankruptcy laws are from the Public Law Research Institute (2005); and corruption measures are from the U.S. Department of Justice, Public Integrity Section.

Table 2 compares characteristics of GAAP and non-GAAP states. The use of debt and the mix of private versus public debt are considered in Panel A. We use one observation (mean or median) per state; therefore, statistical tests compare samples of 13 GAAP state observations (15 GAAP states less two that first mandate GAAP in 2001) with ten observations of unregulated states. The first four rows indicate measures relevant for H1. Differences in measures of the amount of debt issued are not statistically significant, perhaps owing to small sample sizes. Even so, except for the median debt outstanding, the overall use of municipal debt in GAAP states exceeds that of non-GAAP states. The last three rows in Panel A indicate that GAAP states are more likely to issue public debt, although comparisons are not always statistically significant.

Panel B of Table 2 summarizes select features of the sample of 424 public debt issues used to consider whether the GAAP requirement varies cross-sectionally with debt costs (H3). The first row of Panel B indicates that mean true interest cost (TIC)—defined as the rate of return that equates the proceeds of the issue with the present value of interest and principal payments (Benson 1979; Sengupta 1998)—is lower for municipal bond issues in GAAP states than in unregulated states. This comparison is consistent with the proposition that mandating GAAP reporting requirements is associated with lower municipal debt costs. Conclusions are premature, however, as debt costs are influenced by factors other than state

---

[13] Following Figlio and O'Sullivan (2001), we exclude municipalities with population less than 2,000, as Census data are often incomplete.

[14] The 424 observations do not include three observations consistently identified as influential in regressions reported later using Cook's D statistic (Belsley et al. 1980). Results are comparable when the earliest, rather than the most recent, debt issue is selected. We also estimate specifications that omit each state separately, as well as omit the largest two GAAP and non-GAAP states, in order to determine whether results are influenced by the inclusion of any individual state. Results for these procedures are consistent with those reported.

TABLE 2
Comparisons of GAAP and Non-GAAP States

Panel A: Mean (Median) State-Level Municipal Bond Activity, 1995–2002

	All States (n = 23)	GAAP States (n = 13)	Non-GAAP States (n = 10)	t-statistic
Fraction of municipalities that issue debt	0.58	0.59	0.57	0.28
New debt per capita, all municipalities	221.4 (220.6)	237.7 (231.5)	200.2 (200.8)	1.07
New debt per capita, issuing municipalities only	241.6 (245.0)	257.5 (252.7)	220.9 (228.6)	1.04
Outstanding debt per capita	1555.9 (1400.6)	1676.3 (1400.6)	1399.5 (1413.9)	0.97
Fraction of total new debt that is public debt	0.24	0.28	0.18	1.96**
Fraction of municipalities that use private debt only	0.60	0.55	0.64	1.55
Fraction of municipalities that use public debt only	0.06	0.10	0.02	2.08**

Panel B: Mean (Median) Characteristics of Public Debt Issues

	All Issues (n = 424)	GAAP States (n = 212)	Non-GAAP States (n = 212)	t-statistic
TIC	4.61 (4.62)	4.39 (4.46)	4.82 (4.78)	6.00***
Bond ratings:				
Unrated	0.22	0.19	0.24	0.97
Baa	0.01	0.01	0.01	0.48
A	0.05	0.07	0.04	1.25
Aa	0.10	0.17	0.03	4.90***
Aaa	0.61	0.56	0.67	2.57***
Insured	0.58	0.52	0.66	2.44**
Competitively bid	0.82	0.91	0.73	5.25***
Bank qualified	0.69	0.65	0.74	2.12**
GO bond	0.71	0.74	0.69	1.15
Call	0.83	0.75	0.92	4.97***

Panel C: Distributions of State-Imposed Requirements by GAAP Requirements

	All States (n = 23)	GAAP States (n = 13)	Non-GAAP States (n = 10)	Chi-Squared Statistic
Require audits	8	5	3	0.01
Require official statements	10	6	4	0.09
No bankruptcy	4	3	1	0.67

---

*, **, *** Significant at .10, .05, and .01 (two-tailed), respectively.

Variable Definitions:

Bank qualified = dummy variable equal to 1 if the bond issue is bank qualified, and 0 otherwise;
Call = dummy variable equal to 1 if any part of the bond issue is callable, and 0 otherwise;

(*continued on next page*)

TABLE 2 (continued)

Competitively bid = dummy variable equal to 1 if the bond is issued through competitive bid, and 0 if the bond is sold using a negotiated sale;

Fraction of municipalities that issue debt = percentage of municipalities in the state that issued debt from 1995–2002;

Fraction of municipalities that use private debt only = 1995–2002 percentage of municipalities in each state that borrow only from private lenders (banks or institutional lenders such as pension funds or insurance firms);

Fraction of municipalities that use public debt only = 1995–2002 percentage of municipalities in each state that issue only publicly traded debt;

Fraction of total new debt that is public debt = 1995–2002 ratio of the total new public debt to total new debt;

GO bond = dummy variable equal to 1 if the issue is a general obligation bond, and 0 if the issue is a revenue bond;

Insured = dummy variable equal to 1 if any part of the bond issue is insured, and 0 otherwise;

New debt per capita, all municipalities = 1995–2002 mean (median) for all municipalities in the state;

New debt per capita, issuing municipalities only = 1995–2002 mean (median) for all municipalities that issued debt;

No bankruptcy = the state law proscribes municipal bankruptcy; and

Outstanding debt per capita = 1995–2002 mean (median);

Rating = Moody's bond ratings;

Require audits = states that require municipal audits;

Require official statements = states that require official statements when municipal bonds are issued;

TIC = true interest cost, computed as the interest rate that sets the present value of the principal and interest payments equal to the issue proceeds received by the issuer.

regulation (Kessel 1971; Benson 1979; Wilson and Howard 1984; Sengupta 1998; Mansi et al. 2004). In particular, the table shows that bond issues in GAAP states are less likely to be insured or bank-qualified, more likely to be competitively bid, and less likely to have call provisions. Finding greater use of bond insurance for unregulated municipalities is consistent with findings in Gore et al. (2004) that bond insurance and financial disclosure are substitute determinants of municipal debt costs. On the other hand, we find no significant between-sample differences in overall bond ratings (t-statistic = 0.26), although potentially important differences exist in individual rating categories. Finally, differences in the use of general obligation (GO) versus revenue debt are not statistically significant.

Panel C of Table 2 indicates whether GAAP states are more likely than non-GAAP states to require annual independent audits of municipalities, require official statement filings with municipal debt issues, or prohibit municipalities from declaring bankruptcy. These characteristics are used to address concerns about the extent that empirical results reported later can be attributed to state imposed requirements that may be correlated with the GAAP disclosure requirement. Evidence in Panel C indicates that these state requirements are imposed independently of the GAAP requirement. Again, relatively small sample sizes potentially compromise statistical power and, therefore, we incorporate these requirements into the multivariate specifications, which are considered next.

IV. MULTIVARIATE SPECIFICATIONS

Tests of Hypothesis 1

We use the following specification to investigate associations between GAAP disclosure regulation and the use of municipal debt financing.

$$DEBTUSE_i = \alpha_0 + \alpha_1 GAAP\ state_i + \sum \alpha_j CONTROL_{j,i}, \tag{1}$$

where the dependent variable $DEBTUSE_i$ is a measure of municipal i's 1995–2002 use of debt financing (described later), and $CONTROL_{j,i}$ represents j control variables. Parameter estimates α_1 on the dummy variable *GAAP state*, which is set to 1 (0) for bond issues in GAAP (non-GAAP) states, are the focus of the investigation. Specifically, $\alpha_1 > 0$ indicates municipal debt use is greater for GAAP than for non-GAAP states.

Results for Specification (1) are displayed as Table 3.[15] The dependent variable $DEBTUSE_i$ for specifications in the first three columns distinguishes 2,201 municipalities that issue public or private debt during 1995–2002 from 2,754 municipalities that do not issue debt during the period. Parameter estimates in these columns are from a logit specification of the probability of debt issue. Results in the fourth through sixth columns are for an OLS specification of $DEBTUSE_i$ computed as the total new debt per capita issued during the same period. For both dependent variables, we show parameter estimates for subsamples delineated according to size (municipal population). Size partitions consider the possibility that differences in the reporting environment between large and small municipalities cause the consequences of state regulation to manifest differentially between the subsamples. For example, potential lenders may be better informed about large than small municipalities or incentives to comply with GAAP voluntarily may differ between large and small municipalities.

We control for municipal size (log population), 1995–2002 population growth, state regulation of municipal practices other than state GAAP requirements, and individual municipal disclosure (GFOA certificate). The GFOA certificate is awarded to municipal governments that satisfy financial reporting standards set by the GFOA (formerly the Municipal Finance Officers Association). GFOA standards require both GAAP compliance and additional non-GAAP disclosures, although participation in the program is voluntary. Thus, the GFOA variable indicates both GAAP compliance and the extent of voluntary disclosure.

The positive associations between the use of debt financing and the GFOA certificate are consistent with evidence in Evans and Patton (1983) and Gore (2004). Size is associated with greater use of debt financing, while population growth is not significant. Two measures of enforcement (i.e., whether municipal audits are mandated and a high percentage of CPA's on the state auditor's staff), and bankruptcy proscription are associated with significantly less debt financing in some specifications.

Except for the specification in the second column, the estimate α_1 is not statistically significant ($p < 0.10$, two-tailed). These results do not reliably support rejecting H1, although we achieve statistical significance for the logit specification of small municipalities, where potential lenders are likely to be less well informed and/or incentives to disclose voluntarily are less substantial. Therefore, we find no consistent evidence that GAAP disclosure regulation significantly influences debt issuance beyond what is achieved through voluntary disclosure and other state-imposed mechanisms. Failure to detect statistically significant differences indicates that decisions to implement debt-financed municipal projects are more complex than what is suggested by the characterization in Section II. For example, such decisions are certainly influenced by political considerations, as well as by the costs of debt financing.

---

[15] Standard errors for all multivariate specifications are adjusted for within-state clustering using procedures advanced by Rogers (1993) as described in Cohen et al. (2003).

TABLE 3
Associations between the Use of Debt Financing and GAAP Disclosure Requirements

	Logit Model[a]			OLS Model[b]		
	Full Sample (n = 4,955)	**Small[c] (n = 2,478)**	**Large (n = 2,477)**	**Full Sample (n = 4,955)**	**Small (n = 2,478)**	**Large (n = 2,477)**
Intercept	−9.44	−10.96	−8.93	−2.35	−1.57	−3.63
	(−10.60)***	(−7.30)***	(−6.17)***	(−6.13)***	(−2.54)***	(−9.17)***
Ln population (+)	1.12	1.34	1.03	0.31	0.22	0.44
	(11.61)***	(7.77)***	(6.84)***	(6.74)***	(2.75)***	(10.18)***
Population growth (+)	−0.19	−0.46	−0.11	−0.05	−0.05	−0.04
	(−0.85)	(−0.70)	(−1.03)	(−1.88)*	(−0.90)	(−1.43)
GFOA certificate (+)	0.86	0.59	0.93	0.28	0.19	0.28
	(2.97)***	(1.38)	(2.97)***	(3.60)***	(1.49)	(3.88)***
Audits (?)	−1.29	−1.72	−0.89	−0.31	−0.33	−0.28
	(−1.67)*	(−2.03)**	(−1.17)	(−1.09)	(−1.61)	(−0.70)
Official statements (?)	−0.55	−0.75	−0.34	−0.17	−0.13	−0.22
	(−1.04)	(−1.34)	(−0.62)	(−0.89)	(−1.03)	(−0.77)
No bankruptcy (?)	−0.70	−1.02	−0.36	−0.08	−0.07	−0.04
	(−1.34)	(−2.66)***	(−0.55)	(−0.40)	(−0.60)	(−0.14)
Corruption (?)	−0.22	−0.34	−0.13	−0.02	0.03	−0.05
	(−0.55)	(−0.71)	(−0.31)	(−0.17)	(0.27)	(−0.27)
Percentage of CPAs (?)	−0.66	−1.27	−0.15	−0.09	−0.11	−0.04
	(−1.58)	(−2.83)***	(−0.31)	(−0.70)	(−1.43)	(−0.22)
GAAP state (?)	0.26	0.69	−0.08	0.18	0.20	0.15
	(0.72)	(1.81)*	(−0.21)	(1.27)	(1.60)	(0.80)
Adjusted R^2				0.09	0.01	0.15
Model likelihood ratio	1234.68	123.74	447.52			

---

*, **, *** Significant at .10, .05, and .01 (two-tailed), respectively.

[a] Entries in the first three columns are estimates for a logit specification of the probability of using debt financing (public or private) during 1995–2002.

[b] Entries in columns four through six are for an OLS specification of the 1995–2002 amount of new debt per capita.

[c] "Small" and "Large" partitions are at the sample median population.

Variable Definitions:

Ln population = log of mean 1995–2002 population;
Population growth = change in population over the preceding five years;
GFOA certificate = dummy variable equal to 1 if the GFOA certificate is awarded in any of the prior four years;
Audits = dummy variable equal to 1 if the state requires municipal audits;
Official statements = dummy variable equal to 1 if the state mandates official statements;
No bankruptcy = dummy variable equal to 1 if state law proscribes municipal bankruptcy;
Corruption = dummy variable equal to 1 if the ratio of the number of indicted state and local government officials deflated by total state population exceeds the sample median;
Percentage of CPAs = dummy variable equal to 1 if the fraction of CPAs on the state auditor's staff are above the sample median; and
GAAP state = dummy variable equal to 1 if the issuing municipality is in a GAAP state, and 0 if the municipality is in an unregulated state.

Standard errors are adjusted for clustering within states using procedures advanced in Rogers (1993).

Tests of Hypothesis 2

We use the following specification to consider the use of private versus public municipal debt.

$$PRIVDEBT_i = \beta_0 + \beta_1 GAAP\ state_i + \sum \beta_j CONTROL_{j,i}, \tag{2}$$

where $PRIVDEBT_i$ is computed as municipal i's ratio of private debt issued to total debt issued during 1995–2002, and the subscript j indicates a vector of municipal-specific control variables that potentially explain variation in the use of private versus public debt. The dependent variable $PRIVDEBT_i$ is obtained by comparing municipal i's total debt issued from the U.S. Census data with the total public debt issues from the SDC data. We compute the proportion of private debt as the complement of the proportion of publicly issued debt (Krishnaswami et al. 1999; Dhaliwal et al. 2005).[16] Parameter estimates β_1 on the *GAAP state* dummy variable indicate tests of the null H2. Specifically, rejecting the null in favor of $\beta_1 < 0$ supports the proposition that municipalities in GAAP states are more likely to issue public debt than municipalities in states where municipal accounting practices are not regulated.

Results for Specification (2) are in Table 4. Estimates for the full sample are in the first column, and results for small and large municipalities are in the second and third columns, respectively. Control variables—beyond those discussed previously in the analysis of the use of debt financing—are included based on arguments in Dhaliwal et al. (2005; hereafter, DKP), who consider the use of private versus public debt in the corporate context. More specifically, DKP distinguish issues with BBB or higher bond ratings, arguing that firms with lower bond ratings are more inclined to use private debt. The distribution of municipal bond ratings is skewed high relative to corporate ratings and, thus, we distinguish issues with Aaa ratings from the others. DKP also posit that larger debt issues are more likely to be issued publicly than privately. The absolute sizes of municipal debt issues during 1995–2002 are highly collinear with population (see evidence in Table 3) and, therefore, we use the mean debt per capita issued from 1995–2002. DKP include return on assets as a financial performance measure, positing that performance varies directly with the use of public debt. The notion of accounting income lacks substantive interpretation for municipalities; therefore, to consider financial performance, we distinguish municipalities that experience at least one deficit during 1995–2002. Finally, to consider voluntary financial disclosures, we include the fraction of observations during 1995–2002 that earn the GFOA certificate. Notwithstanding these differences in measures, results in Table 4 are consistent with the DKP characterization. In particular, we find negative associations for mean municipal debt per capita, the Aaa bond rating, and voluntary disclosure, and a positive association for the deficit dummy variable.

Turning to the primary focus of the analysis, parameter estimates β_1 on the *GAAP state* variable in the last row of parameter estimates in Table 4, are consistently negative. Moreover, estimates are robust to considering other state-imposed statutory controls for municipalities. As examples, results are comparable regardless of whether state law requires independent audits, official statement disclosure, proscribes municipal bankruptcy, or whether the state has a high corruption index or a high percentage of CPAs on the state auditors' staff. Thus, the null H2 is rejected in favor of the alternative that municipalities

---

[16] Consistent with Dhaliwal et al. (2005) we compute the ratio of private to total debt using the entire sample period, as annual financing decisions can temporarily deviate from the true underlying profile. Results for specifications using sample observations for each municipality-year are consistent with those presented, however.

TABLE 4
Associations between GAAP Requirements and the Ratio of Private to Total Debt[a]

Variable (Expected Sign)	Full Sample (n = 2,049)	Small[b] (n = 1,026)	Large (n = 1,023)
Intercept	1.34	1.87	1.26
	(7.19)***	(5.78)***	(4.64)***
Ln population (−)	−0.05	−0.12	−0.04
	(−4.17)***	(−3.16)***	(−1.98)**
Population growth (?)	0.03	0.02	0.04
	(2.40)**	(1.04)	(2.05)**
Aaa bond rating (−)	−0.40	−0.61	−0.39
	(−7.64)***	(−4.63)***	(−5.52)***
Mean debt issued (−)	−0.06	−0.07	−0.03
	(−2.48)***	(−2.30)**	(−1.37)
Deficit (+)	0.03	0.01	0.08
	(1.46)	(0.67)	(1.75)*
GFOA certificate (−)	−0.05	0.01	−0.08
	(−1.87)*	(0.14)	(−2.77)**
Audits (?)	0.10	0.13	0.08
	(1.14)	(1.72)*	(0.80)
Official statements (?)	−0.02	−0.07	0.03
	(−0.33)	(−1.15)	(0.38)
No bankruptcy (?)	−0.03	−0.05	−0.00
	(−0.44)	(−0.85)	(−0.05)
Corruption (?)	0.06	0.08	0.05
	(0.84)	(1.29)	(0.51)
Percentage of CPAs (?)	0.01	0.01	0.02
	(0.09)	(0.20)	(0.14)
GAAP state (−)	−0.13	−0.10	−0.16
	(−2.63)***	(−2.13)**	(−2.72)***
Adjusted R^2	0.20	0.18	0.17

---

*, **, *** Significant at .10, .05, and .01 (two-tailed), respectively.

[a] Entries are parameter estimates for an OLS specification of the ratio of new private debt to total new debt issued during 1995–2002.

[b] "Small" and "Large" partitions are at the sample median population.

Variable Definitions:

Ln population = log of the mean annual municipal population during the period;
Population growth = change in population over the preceding five years;
Aaa bond rating = dummy variable equal to 1 if the municipality's most recent bond rating is Aaa, and 0 otherwise;
Mean debt issued = mean 1995–2002 total new debt per capita;
Deficit = dummy variable equal to 1 if the municipality reports at least one deficit during 1995–2002;
GFOA certificate = fraction of years during 1995–2002 the municipality is awarded the GFOA certificate;
Audits = dummy variable equal to 1 if the state requires municipal audits;
Official statements = dummy variable equal to 1 if the state mandates official statements;
No bankruptcy = dummy variable equal to 1 if the state does not allow municipal bankruptcy;
Corruption = dummy variable equal to 1 if the ratio of the number of indicted state and local government officials deflated by total state population exceeds the sample median;
Percentage of CPAs = dummy variable equal to 1 if the percentage of CPAs on the state auditor's staff exceeds the sample median; and
GAAP state = dummy variable equal to 1 for municipalities in GAAP states, and 0 for municipalities in unregulated states.

Predicted signs are shown parenthetically.
Standard errors are adjusted for clustering within states using procedures advanced in Rogers (1993).

in GAAP states are less likely to place debt issues privately (more likely to issue public debt) than municipalities in unregulated, non-GAAP states. This evidence is consistent with the premise that disclosure regulation lowers public debt costs beyond what is achieved through voluntary disclosure and other state-imposed regulations.

Tests of Hypothesis 3

Cross-sectional regressions of the following general form consider associations between GAAP disclosure regulation and bond yields.

$$TIC_i = \gamma_0 + \gamma_1 GAAP\ state_i + \gamma_2 INDEX_i + \sum \gamma_j CONTROL_{j,i}, \tag{3}$$

where the subscript *i* distinguishes the most recent bond issue for municipality *i* during 1995–2002, and the subscript *j* indicates a vector of control variables that potentially explain variation in true interest costs (*TIC*).

Parameter estimates γ_1 for the dummy variable *GAAP state* are the principal focus of the investigation. In particular, finding $\gamma_1 < 0$ supports rejecting H3 in favor of the alternative hypothesis that interest costs are lower for municipalities in GAAP states than for municipalities in unregulated states.

The variable *INDEX*, which is the prevailing Bond Buyer market yield at the time of the issue, considers interest rate fluctuations.[17] We use indicator variables to distinguish issue-specific qualitative features that we expect influence bond yields (Kessel 1971; Benson 1979; Wilson and Howard 1984; Sengupta 1998; Mansi et al. 2004). In particular, competition among underwriters reduces bond yields and, therefore, we distinguish issues that are competitively bid. General obligation issues, which are guaranteed by the municipal government and typically also by the state government, sell at lower yields than revenue debt issues, which are backed only by specific revenue streams. Similarly, bank-qualified issues—that is, issues where the total debt issued by the municipality in the year is less than $10 million—receive favorable federal income tax treatment and, therefore, typically sell at lower yields. Because lenders require compensation for the related interest rate risk, callable debt typically sells at higher yields. Recall that the GFOA certificate indicates the extent of voluntary disclosure (and GAAP compliance) by the municipality.

Prior studies find that bond yields vary directly, albeit not linearly, with bond maturity. Maturities of municipal issues in unregulated states tend to be greater than maturities in GAAP states (result not reported). Following Gande et al. (1999), we use two dummy variables to distinguish both short-term maturities (< 5 years) and long-term maturities (> 15 years) from issues with five- to 15-year maturities.

Bond ratings control for credit risk. Following Anderson et al. (2003) and Mansi et al. (2004), we convert default ratings into a single control variable. More specifically, we assign a value of 1 if the Moody's bond rating is Aaa, and as the bond rating declines, the ratings variable increases by 1 (Anderson et al. 2003), with the highest number assigned to unrated issues.[18] As with prior analysis, we use a series of dummy variables to investigate how state requirements other than GAAP disclosure regulation vary with *TIC*. In addition, we include marginal personal income tax rates and state gross product per capita, two measures of state economic condition that potentially influence debt costs.[19] Finally, specifications include seven dummy variables to consider mean effects of factors correlated with the year of issue (1995–2002). To ease the presentation, estimates are not tabulated for year indicator variables.

---

[17] Following Kessel (1971), we also consider specifications (not reported) of risk premia *(TIC-INDEX)*, or the true interest cost net of the appropriate Bond Buyer index. Results using this measure are comparable qualitatively to reported results. Results are also comparable when the dependent variable is computed as net interest costs, a measure that ignores the time value of money, but is commonly used in early studies of municipal borrowing costs.

Specification (3) is applied to the sample of 424 municipal debt issues in Table 5. Consistent with prior literature, we find that most control variables that consider characteristics of the debt issue are statistically significant in predicted directions. On the other hand, with the exception of the *GAAP state* variable, most of the state-level control variables are not significant. Estimates on the parameter γ_1 for the entire sample, displayed in the first column, indicate that mean true interest costs in GAAP states are 16 basis points lower than in unregulated disclosure states. For partitions established by comparing municipal population with the sample median, displayed in the second and third columns, the GAAP-versus-unregulated state mean difference in true interest costs (parameter estimates −0.25 versus −0.14) indicates that the benefit to the GAAP requirement is approximately 11 basis points greater for small municipalities than for large municipalities. This difference, which is statistically significant (two-tailed $p < 0.05$), supports a characterization where greater benefits accrue to small municipalities where lenders are less well informed or where federal regulation is less likely to apply. Overall, we can confidently reject the null H3 in favor of the alternative hypothesis that borrowing costs are lower for GAAP states than for non-GAAP states. Such evidence is consistent with the premise that state-imposed GAAP disclosure regulation reduces municipal debt costs incrementally beyond what is achieved through voluntary municipal disclosure and other state-imposed regulations. Further, the results indicate that state-level GAAP regulations are effective despite the recent federal regulations that encourage voluntary disclosure in unregulated states.

Evidence Regarding the Implementation of GAAP Regulation

The preceding analyses use cross-sectional comparisons to consider associations between GAAP disclosure regulation and municipal debt costs. Although we consider measures that control for the influence of other state-imposed statutory mechanisms, we are concerned about whether results can be attributed to unidentified state-level factors not considered in the specifications. To address such concerns, we consider municipal debt issues in Montana and Nevada, states that first imposed GAAP accounting regulations on 7/1/01 and 6/27/01, respectively.

For these analyses, we use the SDC data to construct a sample of 33 municipal debt issues—11 in Montana and 22 in Nevada—during the calendar year preceding the state GAAP requirement (2000), the year that the requirement is imposed (2001), and the year following the requirement (2002). Note that we execute the analysis using relatively long

---

[18] Results are comparable when the log of maturity is used rather than the Gande et al. (1999) classification, and when a series of dummy variables is used to consider bond ratings categories. Results are also comparable when observations are distinguished according to whether the municipality purchases bond insurance. The effects of bond insurance are apparently subsumed by the bond ratings variables, however, as the correlation between the insurance dummy and the bond ratings variable is −0.76.

[19] Results are consistent for specifications that include measures of state disposable income per capita, state population, and the volume of municipal bonds. Note that continuous measures of state characteristics are unique to each state, and thus distinguish state fixed effects. Hence, we construct dummy variables that partition municipalities according to how corresponding state characteristics compare with median state characteristics.

TABLE 5
Associations between True Interest Cost (TIC) and GAAP Disclosure Requirements[a]

Independent Variable (Expected Sign)	Full Sample (n = 424)	Small[b] (n = 212)	Large (n = 212)
Intercept	−1.95	−1.63	−2.18
	(−2.94)***	(−1.83)*	(−3.41)***
Bond Buyer index (+)	1.22	1.14	1.26
	(10.24)***	(6.77)***	(9.50)***
Competitively bid (−)	−0.13	−0.03	−0.16
	(−1.68)*	(−0.23)	(−1.98)**
GO issue (−)	−0.29	−0.19	−0.35
	(−4.90)***	(−1.99)**	(−5.40)***
Bank qualified (−)	−0.12	−0.10	−0.03
	(−1.97)**	(−1.10)	(−0.48)
Call provision (+)	0.34	0.29	0.34
	(4.97)***	(2.53)***	(2.97)***
GFOA certificate (−)	0.04	−0.06	0.04
	(0.99)	(−0.62)	(0.63)
Maturity < five years (−)	−0.61	−0.66	−0.57
	(−2.42)***	(−1.36)	(−2.48)***
Maturity > 15 years (+)	0.58	0.49	0.66
	(12.69)***	(6.21)***	(8.15)***
Ln issue amount (−)	−0.06	−0.04	−0.04
	(−1.74)*	(−0.89)	(−1.53)
Ln population (−)	−0.07	−0.08	−0.07
	(−4.40)***	(−2.07)**	(−3.87)***
Population growth (−)	−0.00	0.01	−0.00
	(−1.26)	(1.19)	(−1.54)
Tax rates (+)	0.04	0.05	0.08
	(0.62)	(0.55)	(1.18)
State gross product per capita (?)	0.12	0.23	−0.03
	(1.85)*	(3.31)***	(−0.38)
Bond rating (−)	0.02	0.02	0.00
	(5.17)***	(5.21)***	(0.72)
Audits (−)	−0.02	−0.01	−0.12
	(−0.26)	(−0.07)	(−1.96)**
Official statements (−)	−0.03	−0.05	0.10
	(−0.55)	(−0.42)	(1.19)
No bankruptcy (−)	−0.17	−0.17	−0.05
	(−1.80)*	(−0.91)	(−0.53)
Corruption (?)	−0.01	0.11	−0.13
	(−0.19)	(1.02)	(−2.51)***
Percentage of CPAs (?)	0.03	0.13	−0.02
	(0.54)	(1.01)	(−0.38)
GAAP state (−)	−0.16	−0.25	−0.14
	(−3.31)***	(−3.42)***	(−2.82)***
Adjusted R^2	0.81	0.78	0.81

---

*, **, *** Significant at .10, .05, and .01 (two-tailed), respectively.

[a] Entries are OLS estimates for specifications of true interest cost (TIC), computed as the discount rate that equates the present value of the bond principal and interest payments with proceeds.

(continued on next page)

TABLE 5 (continued)

[b] "Small" and "Large" partitions are at the sample median population.

Variable Definitions:

Bond Buyer index = market yields for municipal general obligation or revenue bonds;
Competitively bid = dummy variable set to 1 for sales through a competitive bidding process, and 0 for negotiated sales;
GO issue = dummy variable set to 1 for general obligation issues, and 0 for revenue issues;
Bank qualified = dummy variable set to 1 if the issue qualifies for preferential tax treatment by bank lenders;
Call provision = dummy variable set to 1 if any part of the issue is callable;
GFOA certificate = dummy variable equal to 1 if the municipality is awarded the GFOA certificate;
Maturity < five years (Maturity > 15 years) = dummy variable set to 1 if the years to maturity are less than 5 (exceeds 15);
Ln issue amount = log of the face value of the issue;
Ln population = log of population;
Population growth = change in population over the preceding five years;
Tax rates = indicator variable set to 1 if the state marginal tax rates are above the sample median;
State gross product per capita = indicator variable equal to 1 if the state gross product per capita exceeds the sample median;
Bond rating = Moody's ratings, where Aaa = 1, and the numerical rating increases by 1 as the bond rating declines;
Audits = dummy variable equal to 1 if the issuing municipality is in a state that mandates audits;
Official statements = dummy variable equal to 1 if the issuing municipality is in a state that mandates official statements;
No bankruptcy = dummy variable equal to 1 if the issuing municipality is in a state that does not allow municipal bankruptcy;
Corruption = dummy variable equal to 1 if the ratio of the number of indicted state and local government officials deflated by total state population exceeds the sample median;
Percentage of CPAs = dummy variable equal to 1 if the percentage of CPAs on the state auditor's staff exceeds the sample median; and
GAAP state = dummy variable equal to 1 for municipalities in GAAP states, and 0 for municipalities in unregulated states.

Predicted signs are shown parenthetically. Specifications include year dummy variables to consider mean effects of unspecified time-variant factors (estimates not reported). Standard errors are adjusted for clustering within states using procedures advanced in Rogers (1993).

windows because municipal debt is issued infrequently, and also because regulation is inevitably anticipated, which undermines our ability to confidently identify a unique event date. We use two sets of dummy variables to partition the sample observations according to how bond issue dates compare with effective dates of the state-imposed GAAP regulation.

First, we consider specifications of TIC that include a dummy variable equal to 1 for issues during the one-and-one-half-year period following the effective date of the regulation. Negative parameter estimates on this variable indicate lower mean TIC following the effective date. Second, we consider specifications that include two dummy variables that classify debt issues according to calendar year. In particular, we set a dummy variable equal to 1 for 2001 issues (set to 0 otherwise), and another dummy variable equal to 1 for 2002 issues (0 otherwise). Negative parameter estimates on the first dummy variable indicate lower mean TIC for issues during 2001, the calendar year that the requirement is imposed, than for issues during 2000, the year prior to the disclosure regulation. Negative estimates on the second dummy variable indicate lower mean TIC for issues during 2002, the calendar year after the requirement is imposed than for issues during 2000. For either classification

P 31443

scheme, negative estimates on dummy variables are consistent with the premise that the implementation of GAAP disclosure regulation reduces municipal debt costs.

Results for these specifications are in Table 6. Panel A shows results for Montana (first two columns), Nevada (third and fourth columns), and the two states combined (last two columns), where the dependent variable is computed as TIC less the corresponding Bond Buyer index of market yields (BBI). Estimates displayed in the first, third, and fifth columns indicate that mean post-regulation TIC (net of BBI) are significantly lower than mean pre-regulation means. Estimates displayed in the second, fourth, and sixth columns indicate that means are significantly lower in the year following the regulation year than in the year prior to the regulation year. Results for comparisons of the regulation year with contiguous years are not always statistically significant, although all estimates are negative.

Panel B of Table 6, which shows estimates for multivariate specifications of TIC when factors considered in the cross-sectional analysis in Table 5 are included, demonstrates that estimates in Panel A are not attributable to differences in issue-specific factors. Estimates for dummy variables that classify observations according to how the issue date compares with the effective date of the GAAP requirement support results in Panel A. In particular, estimates in the first column indicate that mean TIC during the 18 months after the requirement is significantly lower than mean TIC during the 18 months prior to the regulation, and estimates in the second column indicate that mean TIC is significantly lower during the year following the regulation year.[20]

Other Procedures

Additional procedures validate the robustness of the results. First, we consider that, prior to issuing debt, municipalities choose between public and private debt. This sequencing of the decision process raises concerns about how selection bias influences TIC specifications in Table 5. To consider this possibility, we estimate a first-stage probit specification of the choice to issue public versus private debt, and include the inverse Mills ratio in the TIC specification.[21] We find that the negative estimate for the *GAAP state* variable is robust to the two-step procedure; moreover, the inverse Mills ratio is not consistently significant.

Second, we examine whether the Table 4 and 5 results hold for the subset of municipalities that undergo Single Audits according to the U.S. Census. This procedure has two objectives. To address the issue of whether municipalities voluntarily comply with GAAP but choose not to apply for a GFOA certificate, we use the audit opinion per the Single Audit report, rather than GFOA certification, as a measure of voluntary GAAP compliance. We include an indicator variable equal to 1 if the municipality received an unqualified opinion, indicating material compliance with GAAP, and 0 otherwise.[22] We also examine a subset of municipalities that are subjected to Single Audits to consider whether GAAP

---

[20] Adjusting TIC for BBI (Table 6, Panel A), or including BBI in the multivariate specification (Panel B), addresses concerns about whether interest costs for all states decreased after 2001. Moreover, pooling the Montana and Nevada samples with the GAAP state sample in Table 5 indicates that Montana and Nevada municipalities have significantly lower interest costs than non-GAAP states in the post-adoption period.

[21] Results are comparable for two alternative approaches. First, we construct the probit specification to distinguish *municipalities* that issue only private debt from municipalities that issue only public debt during 1995–2002. Second, we distinguish *municipality-years* where the municipality issues only private or only public debt for a given year.

[22] The single audit opinion is a noisy measure of GAAP compliance. We find a significant positive correlation between this measure and GFOA certification, however. Further, prior studies indicate that the majority of qualified municipal audit opinions indicate GAAP violations—specifically, inconsistent application of GAAP or a lack of appropriate fixed asset disclosures (Wilson and Howard 1984).

TABLE 6
Regression Specifications of True Interest Cost (TIC) and GAAP Disclosure Requirements for States Newly Adopting GAAP

Panel A: Dependent Variable: TIC Less Bond Buyer Index[a]

Event Date:	**Montana**[b] (n = 11)		**Nevada** (n = 22)		**Full Sample** (n = 33)	
Intercept	−0.42 (−4.76)***	−0.26 (−4.71)***	−0.29 (−1.25)	−0.30 (−1.23)	−0.34 (−2.39)***	−0.29 (−1.75)*
Issues post-regulation	−0.23 (−1.77)*		−0.83 (−2.84)***		−0.65 (−3.28)***	
Issues during the regulation year		−0.35 (−3.16)***		−0.12 (−0.45)		−0.24 (−1.37)
Issues year after the regulation year		−0.38 (−6.01)***		−0.92 (−2.67)***		−0.84 (−3.40)***
Adjusted R^2	0.17	0.36	0.21	0.20	0.19	0.22

Panel B: Dependent Variable: TIC; Controlling for Other Variables Associated with TIC[a]

	Full Sample (n = 33)	
Intercept	−0.14 (−0.04)	7.08 (1.36)
Bond Buyer index (+)	1.02 (2.79)***	−0.05 (−0.08)
Competitively bid (−)	−0.26 (−0.93)	−0.44 (−1.81)*
GO issue (−)	−0.07 (−0.37)	−0.39 (−1.27)
Bank qualified (−)	−0.12 (−0.34)	−0.64 (−1.80)*
Call provision (+)	0.21 (1.39)	−0.02 (−0.15)
GFOA certificate (−)	0.09 (0.38)	−0.59 (−2.36)**
Maturity < 5 years (−)	−0.39 (−0.50)	−0.72 (−1.24)
Maturity > 15 years (+)	0.66 (3.96)***	1.02 (5.90)***
Ln issue amount (−)	−0.20 (−2.71)***	−0.31 (−3.73)***
Ln population (−)	−0.03 (−0.18)	0.01 (0.04)
Population growth (−)	0.06 (1.11)	0.11 (2.25)**
Bond rating (−)	−0.01 (−0.42)	−0.02 (−1.32)
Issues post-regulation (−)	−0.61 (−2.35)**	
Issues during the regulation year (−)		−0.61 (−1.60)

(continued on next page)

TABLE 6 (continued)

	Full Sample (n = 33)	
Issues year after the regulation year (−)		−1.62 (−2.82)***
Adjusted R^2	0.63	0.68

*, **, *** Significant at .10, .05, and .01 (two-tailed), respectively.

[a] Entries are OLS estimates for specifications of true interest cost (TIC), computed as the discount rate that equates the present value of the bond principal and interest payments with proceeds. In Panel A, the dependent variable is TIC less the Bond Buyer index of market yields for municipal general obligation or revenue bonds.

[b] "Montana" and "Nevada" indicate bonds issued by municipalities in Montana and Nevada, respectively.

Variable Definitions:

Issues post-regulation = dummy variable set to 1 if the municipal bond is issued within one-and-a-half years after the effective date of regulation, and 0 if issued within one-and-a-half years prior to the effective date of regulation;
Issues during the regulation year = dummy set to 1 if the municipal bond is issued within the year of the effective date of regulation (2001), and 0 otherwise;
Issues year after the regulation year = dummy variable set to 1 if the municipal bond is issued within one year following the effective date of regulation (2002), and 0 otherwise;
Bond Buyer index = market yields for municipal general obligation or revenue bonds;
Competitively bid = dummy variable set to 1 for sales through a competitive bidding process; 0 for negotiated sales;
GO issue = dummy variable set to 1 for general obligation issues, and 0 for revenue issues;
Bank qualified = dummy variable set to 1 if the issue qualifies for preferential tax treatment by bank lenders;
Call provision = dummy variable set to 1 if any part of the issue is callable;
GFOA certificate = dummy variable equal to 1 if the municipality is awarded the GFOA certificate;
Maturity < five years (Maturity > 15 years) = dummy variable set to 1 if the years to maturity are less than 5 (exceeds 15);
Ln issue amount = log of the face value of the issue;
Ln population = log of population;
Population growth = change in population over the preceding five years;
Bond rating = Moody's ratings, where Aaa = 1, and the numerical rating increases by 1 as the bond rating declines.

Predicted signs are shown parenthetically along with independent variables. Standard errors are adjusted for clustering within municipality using procedures advanced in Rogers (1993).

disclosure regulation adds incrementally beyond what is provided by federal statutes. Results using these procedures are consistent with those presented. Both the relative use of private debt is significantly higher ($t = -2.11$; $p < .05$, two-sided) and TIC is significantly lower in GAAP states ($t = -2.08$; $p < 0.05$, two-sided) than in states with unregulated disclosure. The Single Audit opinion indicator variable is insignificant in all specifications, however.

Finally, the analysis focuses on states that unequivocally require GAAP or are unregulated, even though some states, while not requiring GAAP, do impose unique state-specific disclosure requirements. This raises issues about whether interstate disclosure regulation in the form of GAAP disclosure requirements or intrastate standardization of accounting methods matters more in terms of reducing debt costs. Although addressing this issue comprehensively is beyond the objective of the study, we can offer some preliminary insight. To this end, we supplement the sample used to execute the Table 5 analysis by including municipalities in five states that require state-specific disclosure and examine the effect of interstate versus intrastate disclosure regulations. We use an indicator variable to distinguish municipalities in states that require specific non-GAAP accounting and disclosure. We find that, while the *GAAP state* indicator variable continues to be negative and statistically significant (estimate = −0.22; t-statistic = −2.57), mean debt costs are 47 basis points higher for municipalities in states that require state-specific non-GAAP disclosure (estimate = 0.47; t-statistic = 3.05).[23] Such evidence supports the proposition that comparability in interstate, not intrastate, disclosure is associated with lower municipal debt costs.

V. CONCLUDING REMARKS

Changes in the municipal debt environment, including the 1984 establishment of GASB as the authority for setting municipal accounting standards, and the imposition of a variety of new federal and state regulations, are reasons to reconsider the costs and benefits of disclosure regulation in the municipal government context. Overall, we find that municipalities in states that mandate GAAP accounting and reporting are more likely to pursue public debt than private debt financing options, and that municipal debt financing costs are 14 to 25 basis points lower in GAAP states than in states with unregulated disclosure. We find no evidence that lower financing costs lead to an increased use of debt financing in general, but we do find some limited evidence that the use of debt is greater for small municipalities where contracting costs between municipal borrowers and lenders are likely to be high. We further demonstrate that these results are robust to considering other state-imposed requirements that can affect debt costs. Moreover, we find that two states that first require GAAP during the period encompassed by the study realize significantly lower mean municipal debt costs following the effective date of the regulation. Such evidence indicates that the effects of GAAP reporting and other state requirements are independent of, and incremental to, the effect of each other.

Finally, it is perhaps important to emphasize that the costs of GAAP regulation potentially outweigh benefits that accrue in municipal debt markets and, therefore, we cannot conclude that specific practices recommended by the GASB dominate other possible reporting procedures. Similarly, we cannot meaningfully compare the costs of GAAP compliance with other voluntary mechanisms that potentially yield greater reductions in municipal financing costs. Thus, the analysis does not definitively support conclusions regarding the costs of regulation or decisions by the states to mandate GAAP accounting for municipalities. Even so, the results consistently support the efficacy of GAAP disclosure regulation as a determinant of the cost of municipal financing.

REFERENCES

Anderson, R., S. Mansi, and D. Reeb. 2003. Founding family ownership and the agency cost of debt. *Journal of Financial Economics* 68 (2): 263–285.

Belsley, D., E. Kuh, and R. Welsch. 1980. *Regression Diagnostics, Identifying Influential Data and Sources of Collinearity*. New York, NY: Wiley.

---

[23] For a specification of the ratio of private to total debt, which also includes municipalities with state-specific disclosure requirements, the *GAAP state* indicator t-statistic is −2.58, while the state-specific indicator variable is not statistically significant. Results are also consistent with tabulated results for samples that include, and specifications that distinguish, municipalities in all 24 states that neither impose GAAP nor are unregulated (that is, have state-specific, piecemeal, or hybrid regulations). In particular, the GAAP state variable is associated with lower interest costs and lower relative use of private versus public debt when these observations, which are omitted from the primary analysis, are included and distinguished.

Benson, E. 1979. The search for information by underwriters and its impact on municipal interest cost. *Journal of Finance* 34: 871–884.

———, B. Marks, and K. Raman. 1984. State regulation of accounting practices and municipal borrowing costs. *Journal of Accounting and Public Policy* 3 (2): 107–122.

———, ———, and ———. 1991. The effect of voluntary GAAP compliance and financial disclosure on governmental borrowing costs. *Journal of Accounting, Auditing, and Finance*: 303–319.

Botosan, C., and M. Plumlee. 2004. Are information attributes priced? Working paper, University of Utah.

Bowen, R., L. DuCharme, and D. Shores. 1995. Stakeholders' implicit claims and accounting method choice. *Journal of Accounting and Economics* 20 (3): 255–295.

Bushee, B., and C. Leuz. 2005. Economic consequences of SEC disclosure regulation: Evidence from OTCBB. *Journal of Accounting and Economics* 39 (2): 233–264.

Butler, A. 2004. Corruption and municipal finance. Working paper, The University of Texas at Dallas.

Cohen, R., C. Polk, and T. Vuolteenaho. 2003. The value spread. *Journal of Finance* 58 (2): 609–641.

Dhaliwal, D., I. Khurana, and R. Pereira. 2005. Costly public disclosure and the choice between private and public debt. Working paper, The University of Arizona.

Downing, C., and F. Zhang. 2004. Trading activity and price volatility in the municipal bond market. *Journal of Finance* 59 (2): 899–931.

Easterbrook, F., and D. Fischel. 1996. *The Economic Structure of Corporate Law*. Cambridge, MA: Harvard University Press.

Evans, J., and J. Patton. 1983. An economic analysis of participation in the municipal finance officers association certificate of conformance program. *Journal of Accounting and Economics* 5 (2): 151–175.

Fairchild, L., and T. Koch. 1998. The impact of state disclosure requirements on municipal yields. *National Tax Journal* 51 (4): 733–753.

Figlio, D., and A. O'Sullivan. 2001. The local response to tax limitation measures: Do local governments manipulate voters to increase revenues? *Journal of Law & Economics* 14: 233–257.

Gande, A., M. Puri, and A. Saunders. 1999. Bank entry, competition, and the market for corporate securities underwriting. *Journal of Financial Economics* 54: 165–195.

Gellis, A. 1996. Municipal securities market: Same problems—no solutions. *Delaware Journal of Corporate Law* 21.

Gore, A. 2004. The effects of GAAP regulation and bond market interaction on local government disclosure. *Journal of Accounting and Public Policy* 23 (1): 23–52.

———, K. Sachs, and C. Trzcinka. 2004. Financial disclosure and bond insurance. *The Journal of Law & Economics* 47 (1): 275–306.

———, and C. Trzcinka. 2004. Is there a substitute for mandatory disclosure? Working paper, University of Oregon.

Governmental Accounting Standards Board (GASB). 2005. GASB facts. Available at: http://www.gasb.org/facts/index.html.

Harris, L., and M. Piwowar. 2006. Secondary trading costs in the municipal bond market. *Journal of Finance* 61 (3): 1361–1397.

Healy, P., and K. Palepu. 2001. Information asymmetry, corporate disclosure, and the capital markets: A review of the empirical disclosure literature. *Journal of Accounting and Economics* 31: 405–440.

Higginbotham, S. 2001. New Jersey office pushes for municipalities to adopt GAAP. *Bond Buyer* 335 (31067): 6.

Hildreth, B., and C. Zorn. 2005. Evolution of the state and local government municipal debt market over the past quarter century. *Public Budgeting & Finance* 25: 127–153.

Hume, L. 2002. Many issuers fail to meet disclosure obligations. *Bond Buyer* 340 (31410).

Ingram, R., and R. Copeland. 1982. Municipal market measures and reporting practices: An extension. *Journal of Accounting Research* 20 (2): 766–772.

Kessel, R. 1971. A study of the effects of competition in the tax-exempt bond market. *Journal of Political Economy* 79: 706–738.

Krishnaswami, S., P. Spindt, and V. Subramaniam. 1999. Information asymmetry, monitoring, and the placement structure of corporate debt. *Journal of Financial Economics* 51: 407–434.

La Porta, R., F. Lopez-de-Silanes, A. Shleifer, and R. Vishny. 1997. Legal determinants of external finance. *Journal of Finance* 52: 1131–1150.

———, ———, and ———. 1998. Law and finance. *Journal of Political Economy* 106: 1113–1155.

Leftwich, R. 1980. Market failure fallacies and accounting information. *Journal of Accounting and Economics* 5: 3–30.

Mansi, S., W. Maxwell, and D. Miller. 2004. Does auditor quality and tenure matter to investors? Evidence from the bond market. *Journal of Accounting Research* 42 (4): 755–793.

Miller, D., and J. Puthenpurackal. 2002. The costs, wealth effects, and determinants of international capital raising: Evidence from public Yankee bonds. *Journal of Financial Intermediation* 11 (4): 455–486.

National Association of State Auditors, Comptrollers, and Treasurers (NASACT). 1996. *Auditing in the States: A Summary.* Lexington, KY: NASACT.

———. 2000. *Auditing in the States: A Summary.* Lexington, KY: NASACT.

———. 2002. *Auditing in the States: A Summary.* Lexington, KY: NASACT.

Public Law Research Institute. 2005. Municipal bankruptcy: State authorization under the federal bankruptcy code. Available at: http://www.uchastings.edu.

Reck, J., E. Wilson, D. Gotlob, and C. Lawrence. 2004. Governmental capital markets research in accounting: A review, extension, and directions for future research. *Research in Governmental and Nonprofit Accounting* 11: 1–33.

———, and ———. 2006. Information transparency and pricing in the municipal bond secondary market. *Journal of Accounting and Public Policy* 25 (1): 1–31.

Rogers, W. 1993. Regression standard errors in clustered samples. *Stata Technical Bulletin Reprints* 3: 83–94.

Sengupta, P. 1998. Corporate disclosure quality and the cost of debt. *The Accounting Review* 73 (4): 459–474.

Soybel, V. 1992. Municipal financial reporting and the general obligation bond market: New York City, 1965–1975. *Journal of Accounting and Public Policy* 11 (3): 207–229.

Tiebout, C. 1956. A pure theory of local expenditures. *Journal of Political Economy* 64 (5): 416–424.

U.S. Department of Housing and Urban Development. 1984. *Single Audit Act of 1984: Questions and Answers.* Available at: http://www.hudclips.org.

Wilson, E., and T. Howard. 1984. The association between municipal market measures and selected financial reporting practices: Additional evidence. *Journal of Accounting Research* 22 (1): 207–225.

THE ACCOUNTING REVIEW
Vol. 83, No. 3
2008
pp. 593–628

International GAAP Differences: The Impact on Foreign Analysts

Kee-Hong Bae
Queen's University

Hongping Tan
University of Northern British Columbia

Michael Welker
Queen's University

ABSTRACT: This paper investigates the relation between differences in accounting standards across countries and foreign analyst following and forecast accuracy. We develop two measures of differences in generally accepted accounting principles (GAAP) for 1,176 country-pairs. We then examine the impact of these measures of accounting differences on foreign analysts. In so doing, we utilize a unique database that identifies the location of financial analysts around the world, creating a sample that covers 6,888 foreign analysts making a total of 43,968 forecasts for 6,169 firms from 49 countries during 1998–2004. We find that the extent to which GAAP differs between two countries is negatively related to both foreign analyst following and forecast accuracy. Our results suggest that GAAP differences are associated with economic costs for financial analysts.

I. INTRODUCTION

This paper investigates the relation between differences in accounting standards across countries and foreign analyst following and forecast accuracy. Despite years of discussion about the potential costs and benefits of accounting harmonization, there exists little if any evidence about the existence or magnitude of such costs or benefits. We hypothesize that differences in accounting standards impose costs on foreign financial

---

We thank the editor, two anonymous referees, Sudipta Basu, Patricia O'Brien, Alan Richardson, Dan Thornton, Sergei Sarkissian, Utpal Bhattacharya, Gus DeFranco, Bill Scott, and seminar participants at the Queen's School of Business, the Schulich School of Business at York University, the Rotman School of Business at the University of Toronto, the University of Northern British Columbia, the Northern Finance Association 2007 Conference, the Financial Management Association 2007 Conference, and the Canadian Academic Accounting Association 2007 Conference for their helpful comments. We appreciate the excellent research assistance of Shiheng Wang. Professor Bae thanks the Bank of Montreal and Professor Welker thanks KPMG for generous research support. In addition, research support from the Social Science and Humanities Research Council of Canada (grant # 410-03-1046) is gratefully acknowledged. We are grateful to Thomson Financial for access to their Institutional Brokers Estimate System (I/B/E/S), provided as part of a broad academic program to encourage earnings expectations research.

Editor's note: This paper was accepted by Dan Dhaliwal.

Submitted December 2006
Accepted October 2007

analysts, costs that dissuade analysts from following firms from countries with accounting standards different from the standards in their home countries. We develop two measures of differences in accounting standards for 1,176 country-pairs based on an international survey of generally accepted accounting principles (GAAP) in 2001. We find a strong negative relation between GAAP differences and foreign analyst following. This negative relation remains in tests that control for a wide range of country-pair characteristics such as cultural, economic, geographic, legal, and industrial proximity.

Additionally, for those analysts who do choose to follow foreign firms, we hypothesize that differences in accounting standards are negatively related to earnings forecast accuracy. While we find a negative relation between GAAP differences and forecast accuracy, this relation is weaker than the relation between GAAP differences and foreign analyst following. Our evidence that accounting differences impact financial analysts who specialize in collecting, analyzing, and disseminating financial information suggests that there are potential economic costs associated with variation in accounting rules across countries.

Our paper differs from the prior literature in several important ways. First, we do not argue that any one set of accounting rules is better than others and we do not attempt to assess the quality of any given set of accounting rules. Rather, we propose that analysts tend to avoid following firms from countries using accounting rules that are significantly different from the accounting rules used in their home country and issue less accurate forecasts for such companies when they choose to follow them. Second, we make use of a unique dataset that identifies the location of financial analysts around the world. We examine the effects of accounting differences on financial analysts because analysts are among the most important, sophisticated, and visible users of financial statements, and because analysts explicitly forecast accounting numbers produced by GAAP. We focus on analyst location because the location of analysts who follow a particular firm is of practical interest in assessing worldwide investment interest in the firm's equity shares. In addition, analyst location is expected to correlate with analyst educational background and therefore familiarity with domicile country GAAP and lack of familiarity with GAAP in other countries without incurring additional costs to acquire such familiarity.[1] Third, unlike studies that focus on a limited number of countries, our sample covers firms from 49 countries. This sample not only contributes to the power of our tests, but also enhances the generalizability of our results.

Our main results—that the extent to which accounting standards differ across countries is negatively related to foreign analyst following—still hold even after we control for various analyst and firm characteristics along with a variety of country-pair differences that might also impact analyst activities. Our main results for forecast accuracy also generally hold after controlling for various analyst and firm characteristics along with a variety of country-pair differences, but are weaker than the analyst following results. Our analyst following results are robust to a wide variety of modifications in our research design, and are corroborated by time-series analysis on a sample of firms voluntarily adopting International Accounting Standards.

In addition, we address the fact that not all firms follow the accounting standards of their home country and that there are other sources of variation in the implementation of accounting standards at the firm level within countries. First, we identify firms that use either U.S. GAAP or IAS using Worldscope data that indicate the accounting standards followed for each firm-year. These firms are excluded from our tests that look at differences in accounting standards at the country-pair level based on the local standards in the firm's home country versus the analyst's home country. We also control for two widely available proxies for firm-level differences in the implementation of accounting standards, presence of a Big 4 auditor, or a qualified audit opinion. Our results are robust to the inclusion of these control variables.

---

[1] Our assumption that analysts are not familiar with GAAP in other countries and must incur costs to gain expertise in understanding other countries' GAAP appears consistent with the assumption that foreign investors incur such costs in Barth et al. (1999).

II. RELATED LITERATURE AND MOTIVATION

Despite years of debate about the desirability of accounting harmonization, there is little if any empirical evidence of the existence or magnitude of the benefits from or costs imposed by differences in accounting standards around the world.[2] The objective of our study is to provide such evidence. We begin with a brief overview of the accounting harmonization debate followed by a discussion of how our study relates to the existing literature.

The international effort in harmonizing national accounting standards formally began in 1973 with the establishment of the International Accounting Standards Committee (IASC), which aimed to improve and harmonize financial reporting around the world. However, there was limited adoption of International Accounting Standards before the late 1990s. An important success for international accounting harmonization culminated in 2002 when the EU parliament approved the regulation requiring EU-registered companies to adopt IAS by 2005.[3]

The proponents of accounting harmonization argue that harmonization helps reduce information asymmetry, lowers the cost of capital, and increases capital flow across borders. A single global set of standards facilitates cross-border comparisons of financial data. With accounting harmonization, investors and other market participants could better use their current expertise to analyze firms from countries with accounting standards that they are familiar with. The opponents argue that the characteristics of accounting problems, history, culture, and institutional frameworks in a country determine the form and content of accounting standards. Hence, an appropriate set of accounting standards in one country need not be an appropriate set of accounting standards in another country. Harmonization also increases firms' costs in complying with the new standards. Further, even if a single set of global accounting standards were adopted internationally, lax and inconsistent enforcement may result in limited compliance with the standards and diminished comparability (Ball et al. 2003).

Our study is closely related to the research that has investigated how accounting standards or their implementation affect financial analysts internationally. Hope (2003a) shows that analysts forecast more accurately for firms that disclose their accounting method choice.

---

[2] Ball (2006) also notes the lack of evidence on which to assess the advantages and disadvantages of uniform accounting rules.

[3] In April 2001, the International Accounting Standards Board (IASB) succeeded the IASC for setting accounting standards. The prospective widespread adoption of international standards in the EU prompted the IASB to revise the standards that it had inherited from the IASC. In March 2004 the IASB completed its "stable platform" of standards. The new and revised standards now include new IFRS and the IAS, both of which are now designated International Financial Reporting Standards (IFRS). For simplicity and consistency with the past literature, we refer to these standards as IAS throughout this paper.

Hope (2003b) also finds that strong country-level enforcement is associated with greater forecast accuracy. Basu et al. (1998) show that forecast accuracy is lower for countries with less accrual basis accounting, more market-based accounting, and less accounting choice. Based on a sample of 80 non-U.S. firms that adopted IAS by 1993, Ashbaugh and Pincus (2001) find that the extent of difference between local GAAP and IAS is negatively associated with analysts' earnings forecast accuracy prior to adopting IAS. They also find that analysts' forecast accuracy increases after adopting IAS. Guan et al. (2006) find that analysts provide more accurate forecasts for firms from countries with GAAP that are more similar to U.S. GAAP, and that this effect is stronger in weak disclosure countries and for firms followed by few analysts.

Other researchers have investigated how variation in institutional characteristics other than accounting standards impact analysts internationally. Chang et al. (2000) document considerable variation across countries in analysts' activities. They also find that legal origin helps explain the accuracy of analysts, though Ang and Ciccone (2001) reach the opposite conclusion. Barniv et al. (2005) find that analysts with superior ability and resources consistently outperform other analysts in common law countries, where market forces provide incentives for performance, than in civil law countries, where market-based incentives are less effective. Bae et al. (2007) find that local analysts have an information advantage, i.e., forecast more accurately, than foreign analysts. They also find that the extent of local analysts' advantage is greater in low disclosure, high earnings management environments, and that the information advantage of local analysts is related to U.S. home bias, i.e., the extent to which U.S. investors underweight foreign stocks in their portfolio.

There is also a limited literature that investigates the determinants of analyst following in an international setting. Lang et al. (2004) find that analysts are less likely to follow firms with potential incentives to withhold or manipulate information, and this relation is stronger for firms from low shareholder protection countries. Bushman et al. (2005) find that analyst following increases after initial enforcement of insider trading laws. They further find that this increase is concentrated in emerging markets. However, no study has yet investigated how differences in accounting standards affect analyst following.

Our paper also complements recent research that investigates the relation between firms' choice of accounting standards and international investment allocation decisions. In fact, many firms have voluntarily adopted either IAS or U.S. GAAP to better position themselves in the competition for investors' confidence and scarce capital. Bradshaw et al. (2004) find that conformity to U.S. GAAP helps foreign firms attract U.S. institutional investment since U.S. GAAP is more familiar to these U.S. investors, potentially reducing information-processing costs and providing them with what is perceived to be higher quality information. Covrig et al. (2007) examine the holdings of over 25,000 mutual funds from around the world and find that foreign mutual fund holdings are higher for firms adopting IAS, suggesting that familiar or better GAAP increases foreign investor interest. Empirical assessments of the effects of accounting standards on both investors and analysts can provide valuable, complementary insights, and both approaches have strengths and weaknesses. Investment decisions by investors may be of greater economic significance than analysts' forecasts and recommendations. However, many investors (mutual funds) follow investment strategies unrelated to firm fundamentals (e.g., index funds), while financial analysts almost invariably utilize financial statements. In addition, analysts issue earnings forecasts that allow an assessment of how GAAP differences affect predictions of accounting numbers. We also acknowledge that analyst following could be a function of investor demand, so our results could also reflect the role that GAAP differences play in affecting investors as well as analysts.[4]

We acknowledge that there is no theory that allows us to specify and test comprehensive models of analyst behavior with respect to either the decision to follow firms or to issue more or less accurate earnings forecasts for the firms they follow. With respect to the decision to follow firms, we assume that this decision is made based on the costs and benefits of following foreign firms. The benefits might include, for example, broadening the analyst's industry coverage to include significant foreign firms, increased commissions for their brokerage firm that accompany expanded coverage and investment recommendations, or investment banking fees. The costs presumably include primarily the direct costs of acquiring information about a new firm, but could include reputational costs if the analyst does a poor job of forecasting and providing recommendations about the new firm. For example, Kini et al. (2005) find that forecast accuracy is rewarded with more favorable employment in the international analyst market. Our hypothesis is that the costs of following a foreign firm increase with the extent of GAAP differences between the analyst's domicile country and the home country of the foreign firm.

We also acknowledge that the costs and benefits that are weighed when deciding whether to follow a foreign firm need not be identical to the factors that affect forecast accuracy. Our hypothesis is, again, that the costs of providing accurate forecasts increase with the extent of GAAP differences between the analyst's home country and the home country of the foreign firm. It is possible that the decision to follow a firm indicates that the benefits of following the firm outweigh the costs of learning the effects of different GAAP on the accounting representation of firm performance. If so, then analysts who choose to follow foreign firms might incur those costs and successfully understand the effects of different GAAP such that GAAP differences do not affect forecast accuracy for analysts following foreign firms. We expect that such a view of the costs of following firms and providing forecasts is too simplistic—the costs are primarily effort-related and are therefore continuous in nature. While the decision to follow a foreign firm may signal the decision to incur costs to learn the effects of different GAAP, it is possible that the effects of GAAP differences would not be completely overcome by the foreign analyst. It is, however, reasonable to anticipate that GAAP differences have a greater effect on foreign analyst following than on their forecast accuracy.

Our hypotheses, stated below in alternative form, reflect our assumption that the costs of both following foreign firms and forecasting earnings for foreign firms increase as the number of differences between the analyst's home country GAAP and GAAP in the foreign firm's home country increase:

---

[4] To ensure that our results are not entirely attributable to the relation between investor demand and analyst activities, we perform the following additional analysis. To see the incremental effect of GAAP differences beyond the demand-driven effect, we partition our sample observations into four groups by the level of openness in firm countries and by the level of GAAP differences between firm countries and analyst countries. We sort our sample countries into Low and High openness groups by the median of our sample countries' openness measures. The openness measure is computed as the ratio of the market capitalization of the constituent firms comprising the IFC Investable index to those that comprise the IFC Global index for each country. The data are from Bekaert et al. (2005). The openness variable measures the proportion of firms that are investable by foreign investors and can proxy for demand from foreign investors. Presumably, the higher the ratio is, the higher the demand from foreign investors for equity investment and accordingly the more demand for foreign analyst services. We also sort our sample countries into Low and High GAAP difference groups based on our two GAAP difference measures, respectively. We then compute the mean foreign analyst following in each subgroup. For both the more and less open countries, the difference in the mean foreign analyst following between the low and high GAAP difference group is significant at better than the .01 level. These results hold for both GAAP difference measures. These results suggest that foreign investor demand is not driving our results.

H1: Foreign analyst following is negatively related to the number of differences between GAAP in the home country of the covered firm and GAAP in the domicile country of the analyst.

H2: Foreign analysts' forecast accuracy is negatively related to the number of differences between GAAP in the home country of the covered firm and GAAP in the domicile country of the analyst.

III. RESEARCH DESIGN

Analyst Location

We use Nelson's Directories of Investment Research (1999–2005) for information on the location of financial analysts. Nelson's Directories provide information on nearly 1,700 research firms with nearly 40,000 equity analysts covering publicly traded companies around the world during the period of 1999–2005. We obtain the full names of equity analysts and their country locations from annual volumes of Nelson's Directories for the years 1999 to 2005.

First, we match Nelson research firms to I/B/E/S brokers by broker names and by looking for overlap in analysts' names associated with the brokers between the two databases.[5] For the matched brokers, we match the equity analysts from I/B/E/S and Nelson's Directory by their last names and the initials of their first names. Each volume of Nelson's Directory is usually published in January using data as of November of the previous year. Thus, we classify an analyst's location in the year t volume of Nelson's Directory as applicable to year $t-1$ in the I/B/E/S database.[6] Once the analysts' names are matched, we obtain the country locations of the I/B/E/S analysts. This procedure allows us to identify the country location for approximately 60 percent of all analysts included in the I/B/E/S database for our sample firms.[7]

As discussed earlier, we are interested in identifying analyst location because analyst location is economically significant as a measure of international interest in a firm's equity shares and because location is expected to correlate with GAAP knowledge and familiarity. Implicit in our hypothesis development is the assumption that analysts' locations are stable—i.e., that analysts, by and large, work in the country where they have been trained and educated such that they are most familiar with the GAAP in the country where they are currently located. Of course, like any profession, there is some international mobility among financial analysts. When analysts move from one country to another, the current location no longer necessarily accompanies greater familiarity with local GAAP. International mobility among financial analysts therefore should decrease our ability to observe the effects of differences in GAAP on analysts. However, of the 8,504 cases where a financial analyst is in our database for two consecutive years, analysts change their country location in 474 instances, less than 6 percent of those cases, suggesting that while cross-border relocations by financial analysts do occur, they are not common.

---

[5] In 1998 Nelson became part of Thomson Financial. In September 2000 Thomson Financial acquired Primark, the parent company of I/B/E/S International, and I/B/E/S became a part of the research and analytic unit of the broker/fund management group at Thomson Financial.

[6] We obtain the same results when we classify an analyst's location in year t volume of Nelson's Directory as applicable to year t in I/B/E/S database.

[7] We later provide evidence about the outcome of this matching process on a country-by-country basis.

Measures of GAAP Differences across Countries

There exist a handful of studies that develop measures of differences in accounting standards across countries. Hung (2001) uses an accrual index computed by comparing the accounting standards in 11 areas for 21 countries based on the 1993 *International Accounting Summaries* by Coopers & Lybrand. Relying on the same data source, Young and Guenther (2003) develop an index that captures the degree of disclosure required in a country based on 15 separate disclosure items. Ashbaugh and Pincus (2001) tabulate accounting differences across eight disclosure requirements and four measurement rules relative to IAS for 13 countries. Bradshaw et al. (2004) measure conformity to U.S. GAAP at the firm level based on 13 accounting method choices using data from Worldscope.

Our main interest is in the extent to which accounting standards differ between two countries. Measuring differences in accounting standards is extremely difficult due to the wide range and complexity of accounting standards, creating a very long list of points on which different standards might be said to agree or differ. We adopt two approaches in this study. Our first approach involves identifying a list of 21 important accounting rules based on a review of the past literature and relying on a recent, comprehensive survey of GAAP differences. We then identify differences in these 21 accounting rules between each of the country-pairs in our sample. Our second approach is less judgmental. We use the survey data to identify commonly occurring differences in accounting across countries rather than using the past literature and our own judgment to determine which GAAP differences to focus on. Generally, we find similar results across both measures of GAAP differences.

We rely on the recent *GAAP 2001: A Survey of National Accounting Rules Benchmarked Against International Accounting Standards* (Nobes 2001). In this survey, partners in large accountancy firms from more than 60 countries benchmarked the local accounting standards in their country against IAS, focusing their attention on rules in place as of December 31, 2001. The survey contains information on how local GAAP differs from IAS on 80 key accounting issues, issues incorporating recognition, measurement, and disclosure rules. For each country, the survey captures four types of differences from IAS:

(1) absence of recognition and measurement rules that are present in IAS;
(2) absence of disclosure rules that are present in IAS;
(3) inconsistencies between local GAAP and IAS that could lead to differences for many enterprises; and
(4) other issues that could lead to differences between local GAAP and IAS for certain enterprises.

To help better understand these categories, we provide example items from each category. In the first category, many countries do not require accounting for employee benefits, required under IAS No. 19, or accounting for impairment of assets as required under IAS No. 36. In the second category, common disclosures that are called for under IAS but not required under local GAAP include segment reporting, related-party transactions, and cash flow statements. Category 3 captures differences such as the capitalization of research and development costs and differences in aspects of deferred tax accounting that may affect some industries. Finally, category 4 includes, for example, differences in accounting for long-term construction contracts.[8] Based on a review of the items contained in this list and drawing on prior research that has utilized indices of differences in accounting standards across countries, our first measure of GAAP differences across countries is based on a list of 21 key accounting items that meet three criteria: (1) the item is identified as a key accounting item in the past literature, where we primarily rely on the list of items compiled by Comprix et al. (2003), Bradshaw et al. (2004), and Basu et al. (1998); (2) we can determine that countries with GAAP that does not conform to IAS have GAAP that are similar to one another; and (3) the item creates variation across countries, i.e., at least five countries from our sample are different from the sample countries that conform to IAS for that item but are similar to one another.

---

[8] Ding et al. (2007) also use the GAAP 2001 survey and explore the causes and consequences of the absence of IAS rules (categories 1 and 2) and departures from IAS rules (categories 3 and 4).

From this list of 21 key accounting items, we construct a measure of GAAP differences for each country-pair in our sample as follows. First, for each accounting item in the list, countries that do not conform to IAS receive a score of 1 for that accounting item. All other countries receive a score of 0 for that item. Then, for each item in our 21-item list, we compare each pair of countries in our sample. If both countries conform to IAS (i.e., each has a score of 0) for that item, or if both countries differ from IAS (i.e., each has a score of 1) for that item, then that pair of countries is deemed to have similar GAAP for that item and the country-pair is assigned a "GAAP difference" score of 0 for that item. If one country conforms to IAS (i.e., a score of 0 for that item) and the other country does not (i.e., a score of 1 for that item), then the two countries are deemed to have different GAAP for that item and the country-pair is assigned a "GAAP difference" score of 1 for that item. This procedure is repeated for all 21 accounting items on our list and the total "GAAP difference" score is simply the sum of the scores for that country-pair across all 21 items. This score is our first measure of accounting standard differences that theoretically ranges between 0 and 21 and increases as the extent of difference in GAAP between the two countries increases. We denote this measure of GAAP differences *gaapdiff1*. This measure of GAAP differences is more judgmental than the next measure we describe because (1) the list of items we include is based on the past literature, and (2) we utilize the text included in the *GAAP 2001 Survey* to help ensure coding accuracy and to help ensure that all countries coded as different from IAS for each item have similar GAAP for that item.

Table 1 provides both the list of 21 accounting items used in measuring *gaapdiff1* and the individual country scores for each item. The table indicates which countries have been coded 1, indicating that the country differs from IAS for that item. For simplicity in reporting we do not report the country-pair scores, though they can be easily inferred from the country-level data in Table 1. Luxembourg had the most differences from IAS with 18 out of 21, while there were no differences between Singapore and South African GAAP and IAS. The Luxembourg-Singapore and Luxembourg-South Africa country-pairs also had the most different GAAP, with a total of 18 differences in GAAP. In addition to the Singapore-South Africa country-pair, the U.K. and Ireland also had zero GAAP differences (not surprising since Ireland used U.K. GAAP during our sample period).

This measure of differences in accounting standards across countries has some advantages. First, because the measure is constructed based on survey responses from local partners in large accounting firms, we expect little error in identifying similarities and differences between each country's local GAAP and IAS. Second, because the measure includes items identified as key accounting areas in past literature, it should capture meaningful, substantive differences in recognition, measurement, and disclosure rules around the world. The overlap between the items included in this measure and the items included in lists of key accounting differences in the past literature also affords comparability with the past literature.

TABLE 1
The Initial GAAP Differences Measure (*gaapdiff1*)

Panel A: The 21 IAS Items Making Up the Initial GAAP Differences Measure

Item	IAS Rules	Description—Countries Coded 1
1	IAS No. 1.7	Do not require a primary statement of changes in equity
2	IAS No. 12	Do not generally require deferred tax accounting
3	IAS No. 14	Require no or very limited segment reporting
4	IAS No. 17	Require no or very limited capitalization of leases
5	IAS No. 19	Do not have rules for accounting for employee benefit obligations (other than defined contribution plans in some cases)
6	IAS No. 19.52	Do not have rules for accounting for employee benefits other than pensions
7	IAS No. 2.36	Do not require disclosure of FIFO inventory cost when LIFO is used
8	IAS No. 22.56/38.99	Do not require impairment testing of goodwill or other intangibles with lives in excess of 20 years
9	IAS No. 24	Have no or very limited disclosure requirements for related-party transactions
10	IAS No. 32.18/.23	Do not require that companies account for their financial instruments based on substance over form
11	IAS No. 32.77	Do not require the disclosure of the fair value of financial assets and liabilities
12	IAS No. 35	Do not have rules outlining the treatment of discontinued operations
13	IAS No. 36	Do not have rules calling for impairment testing for long-term assets, or impairments are only recorded when deemed permanent
14	IAS No. 37	Do not have specific rules dealing with provisions
15	IAS No. 37.14	Permit establishing provision when there is no obligation
16	IAS No. 37.45	Do not have rules calling for the discounting of provisions
17	IAS No. 38.42	Permit capitalization of research and development costs
18	IAS No. 38.51	Permit capitalization of some other internally generated intangibles (e.g., brands)
19	IAS No. 7	Do not require a statement of cash flows
20	IAS No. 8.6	Permit a broader definition of extraordinary items
21	SIC 12	Do not require the consolidation of special purpose entities

Panel B: Differences from IAS by Country—Used in *gaapdiff1*

Country/Item	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
Argentina		1	1	1	1					1	1	1	1		1	1	1	1		1	1
Australia					1							1		1			1				
Austria	1		1					1	1		1	1	1		1	1			1	1	1
Belgium			1		1			1	1	1	1	1			1	1	1		1	1	1
Brazil			1	1	1					1		1	1		1	1	1		1		1
Canada	1						1	1							1	1					
Chile	1		1		1		1	1		1	1	1	1		1	1				1	1
China		1			1		1			1	1	1				1				1	1
Czech	1		1	1	1				1	1		1	1		1	1	1		1	1	1
Denmark	1		1	1	1			1		1		1	1	1						1	1
Egypt	1	1	1	1		1				1	1				1						1
Estonia	1				1					1	1	1	1								1

(continued on next page)

TABLE 1 (continued)

Country/Item	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
Finland	1		1	1		1	1		1	1	1	1	1		1	1	1		1		1
France	1			1	1		1	1	1	1		1	1		1	1				1	
Germany	1						1	1	1	1	1	1			1	1				1	1
Greece	1	1	1	1		1	1		1	1	1	1			1	1	1	1	1	1	1
Hong Kong					1					1	1										
Hungary	1	1	1		1		1		1	1	1	1	1		1	1				1	
India	1									1	1	1	1		1	1	1				
Indonesia						1	1				1										1
Ireland										1											
Israel			1			1		1				1				1					1
Italy	1		1	1						1	1	1	1		1	1			1	1	1
Japan	1			1	1		1					1	1		1	1				1	
Korea					1					1	1				1	1					1
Luxembourg	1	1	1	1	1		1	1		1	1	1	1		1	1	1	1	1	1	1
Malaysia						1				1	1	1	1	1				1			1
Mexico													1								
The Netherlands	1														1	1				1	
New Zealand					1								1					1			
Norway	1				1			1		1			1		1		1				
Pakistan										1		1	1	1							
Peru					1																
Philippines				1		1		1		1	1		1		1	1	1				1
Poland			1		1		1	1		1	1		1		1	1	1			1	1
Portugal	1	1					1		1	1	1	1	1	1			1	1		1	1
Russia	1	1			1		1		1	1	1	1	1		1	1	1	1	1	1	1
Singapore																					
Slovenia	1	1	1		1		1				1	1			1					1	
South Africa																					
Spain	1		1	1			1	1	1	1	1	1	1		1	1	1		1	1	1
Sweden	1					1		1		1	1	1	1		1	1	1				
Switzerland	1		1			1	1			1	1	1	1		1	1				1	1
Taiwan							1			1		1	1			1					1
Thailand		1			1							1		1							
The U.K.										1											
Turkey	1	1	1	1	1		1		1	1	1	1	1			1	1			1	
The U.S.								1		1					1	1					
Venezuela					1					1	1	1				1					

The IAS numbers used for reference in this table are the primary IAS standard references given in the GAAP 2001 survey used for coding each item. In each case, the textual description of the GAAP difference was also utilized in the coding, so in some cases countries identified as differing from IAS on the indicated standard are not coded 1 because the text description for that country suggested that the difference from IAS GAAP was minor and unlike the other countries coded as 1. In addition, in some instances countries are coded as 1 even if the indicated standard was not mentioned for that country, if the text for a related item suggested that the country's GAAP was in fact different from IAS on the dimension described above. Since we are interested in GAAP differences across countries, *gaapdiff1* is the sum of the number of differences observed in the scores across all 21 items between each country-pair; *gaapdiff1* therefore has a theoretical maximum value of 21 and minimum value of 0.

Our approach also has some limitations. First, even though we identify a relatively comprehensive list of items that meet the criteria we discussed earlier, this measure is obviously incomplete and does not capture all important aspects of differences in accounting standards across countries. Second, the measure is based on rules in effect in 2001. To the extent that accounting rules changed in the 1998–2004 sample period, we measure differences in accounting standards with error. Third, and perhaps most importantly, this list captures differences in accounting standards, not necessarily actual practice, across countries. To the extent that companies within these countries adopt accounting standards other than their home country standards, provide voluntary disclosures that go beyond those specified in local standards, or fail to comply with local standards, we measure with error the extent of accounting differences at the company level.[9]

We utilize an alternative measure of accounting differences to ensure that our results are robust to the measurement of accounting differences. The second measure we construct also employs the same *GAAP 2001 Survey* but implements a more mechanical approach to identifying key accounting areas and coding differences between IAS and local GAAP. Our second approach involves two steps. Each accounting practice covered in the GAAP 2001 survey is coded by the IAS section number, along with a brief description of the nature of how local GAAP and IAS differ for that item. In creating our first measure, *gaapdiff1*, we perform the coding exercise by hand, utilizing both the IAS section numbers and the text description to ensure that our coding scheme captures "similar" differences from IAS across countries.[10] For our second measure, we use an algorithm based solely on the IAS section numbers reported in the *GAAP 2001 Survey* results to (1) identify the important accounting practices to include as individual items in our list, and (2) to code cases where local GAAP does not conform to IAS. First, we identify each IAS section that appeared in the individual country reports for five or more countries included in the survey. This generates a list of 52 distinct IAS sections. Next, for each of these 52 IAS sections we simply code a 1, indicating nonconformity with IAS for each country that includes that particular IAS section number in the country report, without referring to the text to ensure that similar departures from IAS are identified for each country.

Because the types of departures from IAS GAAP under the first two categories of departures from IAS (categories dealing with the absence of recognition, measurement, or disclosure rules) tend to be quite similar across countries while the departures from IAS reported under the last two categories (differences under some circumstances or for some companies) are more subtle differences that are less likely to be consistent across countries, we use only the information appearing under the first two types of departures from IAS for coding the country-level departures from IAS. The construction of the country-pair

---

[9] As reported later, we partially address this point by collecting data about the actual accounting standards adopted at the firm level. In addition, we perform additional analyses (1) using measures of disclosure (based on CIFAR scores) and earnings management (from Leuz et al. 2003) differences across countries and (2) controlling for auditor type and audit opinion at the firm level.

[10] Our *gaapdiff1* coding has two features that help ensure that two countries coded as different from IAS have similar GAAP. First, a number of the accounting rules we examine involve the presence or absence of required disclosures (e.g., statements of cash flow and changes in equity). In addition, we read the textual description of the departure from IAS for each country to ensure that each country had similar rules. In over 80 percent of the departures from IAS we identify, our coding reflects little judgment and is nearly unambiguous. For example, all 19 countries coded as departing from IAS because they do not require the disclosure of FIFO inventory when LIFO is used have statements in the GAAP survey that clearly state that this disclosure is not required. In the remaining cases, we apply some judgment in coding departures from IAS even though the textual descriptions of the departures from IAS are not virtually identical across countries. In most cases, this involves coding a country as departing from IAS when the textual description indicates that the IAS rule is followed in only very limited circumstances in that country.

GAAP differences then mirrors the procedures we follow for *gaapdiff1*. We label this measure of GAAP differences *gaapdiff2*. The Appendix details the accounting items that are included in the construction of *gaapdiff2*.[11]

Our first GAAP difference measure, *gaapdiff1*, covers 21 accounting items for a total of 1,176 country-pairs. This measure ranges from 0 to 18, with a mean of 9. Our second GAAP difference measure, *gaapdiff2*, covers 52 accounting items for a total of 1,176 country-pairs. This measure ranges from 0 to 25, with a mean of 13.1. The Pearson (spearman) correlation between *gaapdiff1* and *gaapdiff2* is 0.57 (0.52), implying that our measures capture similar phenomena, though the correlation is modest enough to suggest that each measure captures different aspects of GAAP differences across countries to some extent.

Our discussion and empirical analysis reflect the assumption that accounting standards are exogenous, i.e., that differences in accounting standards across countries impact foreign analyst behavior, but that foreign analyst activities and their demands for information do not impact differences in accounting standards. We acknowledge that foreign investor and analyst demand for information is certainly one factor that could impact how accounting standards are set at the national level. However, we suspect that this source of demand influencing accounting standards is, at best, a second-order effect relative to within-country issues. Absent fundamental changes such as a country's decision to adopt IAS, accounting standards are relatively stable and do not respond quickly to demand forces such as emerging demands for fewer differences with another country due to increased investor and analyst demand from that country. In addition, given the diversity in accounting standards evident in Table 1, it is not clear how standard setters could respond to calls for reducing accounting differences with other countries when those calls could come from countries with quite different accounting standards. Finally, it would be very difficult to empirically model potential endogeneity of accounting standards in the absence of theoretical guidance to assist in identifying exogenous determinants of accounting standards. For these reasons, we do not attempt to structure our empirical analysis to allow for the potential endogeneity of accounting standards.[12]

To summarize, we utilize two different measures of differences in accounting standards across countries. The variable *gaapdiff1* is a judgmentally prepared list, which identifies 21 key accounting areas from the past literature and uses both IAS section numbers and the accompanying text from the *GAAP 2001 Survey* to identify GAAP differences across countries. For *gaapdiff2* we also use the *GAAP 2001 Survey* to identify accounting differences across countries but adopt a more mechanical approach to identify 52 important areas of accounting and to determine country-level conformity to or departure from IAS for each of these 52 areas. Each measure has advantages and disadvantages, and neither is perfect or complete. Importantly, as we show later, our main results are generally quite consistent across both measures, enhancing our confidence in the results.

Dependent and Control Variables

The main purpose of this paper is to examine whether accounting differences across countries are related to foreign analyst following and forecast accuracy. We compute our measure of foreign analyst following, for_anal_{ja}, as the average number of foreign analysts per year from country a who forecast annual earnings for firm j during the seven-year period of 1998–2004. We use the average number of analysts from each foreign country rather than our full panel data because the number of analysts following a company cannot be considered independent across time.[13] For a particular firm, if GAAP differences matter, then there would be more analysts covering the firm from the countries that have fewer GAAP differences from the firm's country than from the countries that have more GAAP differences. To measure forecast accuracy we use the price-scaled absolute difference between an earnings forecast and the actual earnings for firm j at time t. We deflate by stock price to facilitate comparison across firms. The stock prices $P_{j,T-1}$ are the latest available monthly stock price in the previous year from the I/B/E/S summary file.[14] We apply the same price $P_{j,T-1}$ to all of the forecasts made for firm j for the fiscal year T.[15]

---

[11] Because of the large number of IAS sections and countries involved, we do not report the country by country coding of departures from IAS for this list. The country-level departures from IAS used in the construction of *gaapdiff2* are available from the authors upon request.

[12] We do report the results of time-series analysis of firms voluntarily adopting International Accounting Standards later in the paper. This analysis is less subject to potential endogeneity concerns than our cross-sectional tests.

Previous research suggests various factors that might affect analyst following across firms, such as firm size, ADR status, and forecast difficulty.[16] Previous studies have also documented that forecast accuracy depends on forecast horizon, analysts' forecast ability, resources, and forecast complexity (Ang and Ciccone 2001; Clement 1999). Thus, we incorporate various proxies for these factors as control variables. We use analysts' general forecast experience and firm-specific experience to proxy for their ability, and use the number of analysts working with their research firms to proxy for the resources available to the analysts.

We also draw on past literature to identify a wide range of variables that may differ across countries and influence analyst activities. We control for these variables because they are likely to be correlated with GAAP differences. Including these variables is important to assess the incremental effect of GAAP differences on analysts after controlling for other important differences between countries. Our comprehensive list of country-pair control variables is drawn primarily from Sarkissian and Schill (2004), who construct a comprehensive set of country-pair level variables in their analysis of firms' cross-listing decisions. The list of control variables they construct captures virtually all of the important country-level characteristics that have appeared in the prior international finance and accounting literatures. We provide estimates with and without these country controls to assess the impact of including these control variables on our results. The details of variables used in the study including definitions, measurement, and data sources are described in Table 2. We discuss the variables in more detail as we discuss our results.

---

[13] We later report sensitivity tests using alternative measure, *for_anal/total*, which scales the number of analysts as defined above by the average annual total number of foreign analysts following firm j during 1998–2004. We obtain similar results when we restrict our sample to year 2001, the year GAAP differences are measured, instead of using the averages for both dependent and independent variables during the sample period.

[14] We obtain similar results when we use monthly stock prices from Compustat.

[15] We later report results of sensitivity analysis using a second measure of forecast accuracy, the absolute forecast error scaled by actual earnings disclosed in I/B/E/S.

[16] Our primary reference for control variables in the analyst following model is Lang et al. (2003), which examines analyst following in an international context focused on equity cross-listing. We do not include the earnings surprise variable in their model because they find it is not related to analyst following. We substituted forecast diversity for the returns-earnings correlation and the standard deviation of returns (all can be viewed as measure of uncertainty and/or information asymmetry which complicate the forecasting process) because we lose around 1/3 of our sample if we use the later two variables (which require return data). However, adding return standard deviation and the return-earnings correlation to our analyst following regressions does not qualitatively alter our conclusions. Both GAAP difference measures remain significant in the analyst following regressions. The coefficients on return standard deviation and the return-earnings correlation have signs consistent with those reported in Lang et al. (2003).

TABLE 2
Variable Descriptions

Panel A: Dependent Variables

for_anal_{ja}	Average annual number of analysts who follow firm *j* from another country during 1998–2004. It is 0 if there is no analyst from country *a* (identified by various volumes of Nelson's directory) following firm *j*. We multiply by 100 for expositional purposes only.
$afep_{ijt}$	Absolute forecast errors scaled by the most recent stock prices in the previous year. It is computed as the price-scaled absolute difference between an earnings forecast by analyst *i* and the actual earnings for firm *j* at time *t*. To facilitate exposition we multiply this variable by −100 so that larger values represent more accurate forecasts.

Panel B: Independent Variables

$gaapdiff1_{fa}$	Total number of GAAP differences between country *f* and country *a* based on GAAP 2001 survey as described in Panel A of Table 1.
$gaapdiff2_{fa}$	Total number of GAAP differences between country *f* and country *a* based on GAAP 2001 survey as described in the Appendix.

Panel C: Control Variables

$horizon_{ijt}$	Forecast age in years between the forecast date *t* and the corresponding I/B/E/S report date of the actual earnings.
mv_{jY}	Market capitalization in U.S. $ billions for firm *j* in year *Y*. It is the product of the number of shares outstanding and monthly stock price from I/B/E/S summary files.
$firmex_{ijt}$	Analyst firm-specific experience, defined as the time interval in years between analyst *i*'s first forecast for a particular firm *j* in I/B/E/S database and his forecast at time *t* for firm *j*.
$genex_{it}$	Analyst general experience, defined as the time interval in years between analyst *i*'s first forecast in the I/B/E/S database and his current forecast at time *t*.
$brsize_{iY}$	Brokerage size, defined as the number of analysts working for the I/B/E/S brokerage that analyst *i* is associated with in year *Y*.
$nfirm_{iY}$	Number of firms analyst *i* covers in year *Y* in the I/B/E/S database.
$nctry_{iY}$	Number of countries analyst *i* covers in year *Y* in the I/B/E/S database.
$diversity_{jY}$	Diversity of analysts' forecasts for a firm in year *Y*. It is computed as the ratio of standard deviation of forecasts for firm *j* in year *Y* in the I/B/E/S database to the product between the absolute value of consensus forecast and the square root of the number of analyst following firm *j* in year *Y*.
adr_{jt}	Dummy variable indicating whether firm *j* has an ADR program in place at time *t*. For the sample of analyst following, it takes the value of 1 if firm *j* has ever had an ADR program since 1990, and 0 otherwise. The data are from the website of the Bank of New York (2001).
$dist_{fa}$	Geographic distance in thousands of kilometers between firm's country and analyst's country based on latitudes and longitudes of the important cities. The data source is CEPII at http://www.cepii.fr/anglaisgraph/bdd/distances.htm.
$cult_{fa}$	Dummy variable equals 1 if firm's country and analyst's country share either a common language or were historically part of the same colonial empire. The data source is CEPII at http://www.cepii.fr/anglaisgraph/bdd/distances.htm.
$econ\_prox_{fa}$	Percentage of firm-country's exports going to analyst's country for the year 2001. The data are from the TradeAnalyzer database.

(*continued on next page*)

TABLE 2 (continued)

cap_corr_{fa}	Proxy for correlation of industry structure between firm's country and analyst's country. First we calculate the average weight of each one-digit industry (up to ten industries in total) for each country during 1998–2004 by market capitalization. We then calculate the Pearson correlation of the industry weights for each country-pair. The data source is Worldscope.
$stkturn_{fa}$	Ratio of firm-country's stock turnover rate to analyst-country's stock turnover rate based on data from World Development Indicators.
gdp_{fa}	Ratio of firm-country's GDP to analyst-country's GDP based on data from World Development Indicators.
$capgdp_{fa}$	Ratio of firm-country's GDP-scaled equity market size to analyst-country's GDP-scaled equity market size based on data from World Development Indicators.
$legorig_{fa}$	It is the difference between firm's country and analyst's country based on a dummy variable that equals 1 if a country is of English common law origin, and 0 otherwise. The raw data are from La Porta et al. (1998).
$antidir_{fa}$	It is the difference between firm's country and analyst's country in the Anti-Director Rights Index. The original index ranges from 0 to 5, with higher scores for stronger investor rights. The raw data are from La Porta et al. (1998).
$crosslist_{fa}$	The number of firms from firm's country that cross-listed in analyst's country scaled by the number of total domestic firms in firm's country. The cross-listing data are from Sarkissian and Schill (2004) and the number of domestic firms is from World Development Indicators.
$tot\_for\_anal_{a}$	The average total number of analysts from the analyst's country who follow at least one foreign firm during the 1998–2004 sample period.
$cifar_{fa}$	The difference between firm's country and analyst's country in the CIFAR index of financial disclosure. The data are from Bushman et al. (2004).
$earn\_mgmt_{fa}$	The difference between firm's country and analyst's country in the index of earnings management based on data from Leuz et al. (2003).
$big4_{j}$	Dummy variable equals 1 if a firm is audited by one of the Big 4 accounting firms based on data from Worldscope, and 0 otherwise. This is a static data item in Worldscope so it is measured as of the latest fiscal year reported in the database for firm j.
$qualified_{jY}$	Dummy variable equals 1 if a firm receives a qualified opinion in year Y based on data from Worldscope, and 0 otherwise.

In this table i = analyst, j = firm, t = time of the forecast, f = country location of the firm, a = country location of the analyst, and Y = year.

IV. SAMPLE AND DESCRIPTIVE STATISTICS

Sample

We impose the following criteria to yield our final sample:

1. Country-level GAAP differences relative to IAS are available in the *GAAP 2001 Survey*.
2. The location of equity analysts can be identified. Because we are interested in how GAAP differences affect foreign analysts, we exclude local analysts from our final sample. Therefore, a firm must be followed by at least one foreign analyst to be included in our sample.

(3) The annual earnings forecasts are restricted to the latest forecast for each firm/year/analyst with a minimum of 30 days before the I/B/E/S report dates but no longer than one year before the fiscal year-end.[17]
(4) Actual earnings are available from the I/B/E/S actual data file.

The intersection of I/B/E/S and the GAAP 2001 survey yields 1,176 distinct country-pairs. Because our tests examine the effects of GAAP differences between the firm's local GAAP and the GAAP of the foreign analysts' home country, we exclude firm-years from our primary sample when companies do not use local standards based on Worldscope yearly data. Worldscope data on the accounting standards used by the firm are available for approximately one-half of our sample. In our primary tests we retain firms that do not have data on the accounting standards used by the firms, reflecting our assumption that smaller firms without standards data are very likely to be local standard users. Our sample for forecast accuracy covers 6,888 analysts making a total of 43,968 forecasts for 6,169 firms from 49 countries during 1998–2004. We eliminate the extreme 1 percent of our sample of forecast accuracy to eliminate the effects of possible data errors.

Our sample for analyst following is constructed as follows. We compute the average of the annual number of foreign analysts that follow each firm from each of the 48 foreign countries during 1998–2004. Foreign analyst following is 0 if there is no analyst following for a firm from a particular country over our sample period.[18]

Descriptive Statistics

Table 3, Panel A shows the distribution of firms, analysts, and firm market capitalization by country. The analyst match rates reported in this panel show the proportion of I/B/E/S analysts following firms in each country for whom we are able to identify location using Nelson's Directories. This match rate varies around the mean of about 60 percent, ranging from a low of 40 percent (India and Malaysia) to a high of 75 percent (Canada). One concern with the variation in match rates is that we could systematically identify more foreign analysts for larger and more developed countries than for less developed countries, and this bias could be related to our measures of GAAP differences. While we cannot entirely dismiss biases in our sample construction as a contributor to our results, the fact that our match rates are 40 percent or more for all countries and that most country match rates fall reasonably close to the sample mean of 60 percent suggests that biases in sample construction are unlikely to have a significant effect on our results.[19] Sixteen out of a total of 49 countries have more than 100 firms in our final sample. Thirty-four countries have at least 40 firms. There are 27 countries with 100 or more foreign analysts covering firms in their market. Forty-two countries have at least 40 foreign analysts. The analysts in our sample come from 46 sample countries; the U.S. and the U.K. have the largest analyst

---

[17] Our main results remain qualitatively the same when we restrict the annual forecasts to at least five days before the I/B/E/S report dates.

[18] We also repeated our test on a smaller sample of firms that are followed by at least five foreign analysts to see if very low foreign coverage is driving our results. Our results are qualitatively unchanged for this sample. In addition, our full sample includes firms with an ADR program that may be reconciling their accounting numbers to U.S. GAAP. If we drop all firms without accounting standard data and all firms with an ADR (and omit the ADR dummy variable), our results for analyst following remain qualitatively unchanged, but the negative relation between *gaapdiff2* and forecast accuracy becomes insignificant in this smaller sample.

[19] As we discuss in more detail later, we control for the effects of variation in the number of analysts we identify across countries by including the total number of analysts residing in the analyst's country as a control variable in our analysis of analyst following. We also report the results of several additional specifications that deal with variation in the number of analysts across countries by altering how we construct our sample later in the paper. Our main inferences are unaffected in these tests.

TABLE 3
Summary Statistics

Panel A: Number of Covered Firms, Analysts Covering Firms from the Country, Analysts Domiciled in the Country, Number of Forecasts for Covered Firms, and Mean Market Capitalization by Country[a]

Country	Analyst Match Rate	No. of Firms	No. of Analysts Covering Firms	No. of Analysts Domiciled	No. of Forecasts for Covered Firms	Mean Market Cap (U.S.$ billion)
Argentina	0.50	34	101	21	301	1.290
Australia	0.47	151	61	130	271	1.951
Austria	0.57	26	99	6	142	0.674
Belgium	0.65	75	383	61	802	2.669
Brazil	0.43	95	164	59	886	3.475
Canada	0.75	519	1,658	266	5,715	2.276
Chile	0.56	32	107	18	338	1.237
China	0.50	103	188	52	536	1.115
Czech	0.51	21	42	3	97	0.294
Denmark	0.58	62	323	50	716	1.338
Egypt	0.74	14	26	1	77	0.591
Estonia	0.49	3	11	0	12	0.694
Finland	0.60	79	331	24	1,010	1.127
France	0.61	261	1,272	502	4,083	4.189
Germany	0.59	218	807	291	2,434	4.177
Greece	0.59	46	78	1	268	1.095
Hong Kong	0.46	182	212	519	654	2.980
Hungary	0.54	15	34	4	130	0.158
India	0.40	92	75	20	373	1.363
Indonesia	0.44	58	82	18	292	6.531
Ireland	0.67	35	180	38	413	2.159
Israel	0.46	17	36	17	84	1.021
Italy	0.61	141	736	24	2,187	4.075
Japan	0.44	225	45	70	344	7.385
Korea	0.51	71	66	51	177	2.704
Luxembourg	0.69	2	11	0	12	1.380
Malaysia	0.40	114	174	32	714	0.830
Mexico	0.53	77	175	48	1,194	1.496
The Netherlands	0.61	114	867	116	2,368	4.711
New Zealand	0.52	58	69	17	180	1.249
Norway	0.58	62	269	45	597	1.303
Pakistan	0.58	11	6	3	17	0.312
Peru	0.53	20	53	6	125	0.387
Philippines	0.49	41	101	16	273	0.613
Poland	0.57	46	64	3	357	0.636
Portugal	0.57	42	208	19	499	1.475
Russia	0.64	20	41	8	106	3.269
Singapore	0.44	78	156	205	406	1.664

(continued on next page)

TABLE 3 (continued)

Country	Analyst Match Rate	No. of Firms	No. of Analysts Covering Firms	No. of Analysts Domiciled	No. of Forecasts for Covered Firms	Mean Market Cap (U.S.$ billion)
Slovenia	0.55	2	2	0	4	0.393
South Africa	0.57	154	77	51	404	1.264
Spain	0.57	106	634	104	1,877	3.815
Sweden	0.62	141	625	153	2,087	1.867
Switzerland	0.62	58	341	57	719	4.830
Taiwan	0.49	127	162	49	711	2.234
Thailand	0.46	77	150	24	540	0.582
The U.K.	0.61	483	601	2,326	1,958	5.235
Turkey	0.42	73	48	1	302	0.353
The U.S.	0.69	1,680	1,386	1,759	6,109	8.479
Venezuela	0.53	8	43	5	67	0.595

Panel B: Summary Statistics for the Sample of Analyst Forecast Accuracy[b]

Variable	Mean	Median	Std. Dev.	Minimum	Maximum
afep	−2.039	−0.966	3.661	−45.040	0.000
horizon	0.490	0.436	0.246	0.088	2.201
firmex	0.983	0.585	1.489	0.000	20.327
genex	3.446	2.287	3.697	0.000	22.392
nfirm	10.352	9.000	7.560	1.000	140.000
nctry	2.712	2.000	1.555	1.000	13.000
brsize	131.487	103.000	115.837	1.000	444.000

Panel C: Summary Statistics for the Sample of Analyst Following[c]

Variable	Mean	Median	Std. Dev.	Minimum	Maximum
sum_for_anal	2.517	1.333	2.824	1.000	28.343
for_anal	0.052	0.028	0.059	0.021	0.590
adr	0.126	0.000	0.332	0.000	1.000
mv	4.508	0.754	14.425	0.001	345.502

---

[a] The final sample includes 6,888 distinct analysts and 6,169 firms from 49 countries during 1998–2004. The annual earnings forecasts are restricted to the latest ones for each firm/year/analyst with a minimum of 30 days before the I/B/E/S report dates. The number of analysts covering firms and the number of analysts domiciled reported in this table are larger than the 6,888 distinct analysts in our sample because analysts may cover firms from more than one foreign country and in some cases analysts switch country of domicile.

[b] For *afep*, *horizon*, *firmex*, and *genex*, we first compute the median of the variables for each distinct analyst, and then compute the summary statistics across analysts for each variable. For the remaining variables, the descriptive statistics are based on the observations for each analyst year.

[c] In this panel, *sum_for_anal* is the average number of foreign analyst following during 1998–2004 from all of the other countries and *for_anal* is the average number of foreign analyst following during 1998–2004 from each of the other countries. For *adr* and *mv* we first compute the mean for each firm, and then compute the summary statistics across firms.

populations in our sample. The covered firms are the smallest in Hungary and the Czech Republic, with an average market capitalization of only U.S.$158 and U.S.$294 million, respectively. The sample firms in the U.S. and Japan are the largest, with an average market capitalization of more than U.S.$7 billion.

Table 3, Panel B shows the summary statistics for the sample of analysts' forecasts. The mean absolute forecast error scaled by price is around 2 percent. The mean forecast horizon is about 0.5 year. On average the analysts in our sample have followed a firm for about one year and been in their business for more than three years. These analysts cover an average of ten firms, 2.2 I/B/E/S industry groups, and 2.7 countries.[20] They work in brokerage houses employing more than 130 analysts on average.

Finally, Table 3, Panel C shows the summary statistics for the sample of analyst following. We calculate the number of foreign analysts from all foreign countries for each firm in this panel. The mean number of foreign analysts for each firm from all of the other countries is 2.5, while the median is 1.33.[21] An ADR program is in place at some point since 1990 for 12.6 percent of our sample firms. There is substantial variation in firm size across our sample firms.

V. EMPIRICAL RESULTS

Foreign Analyst Following and GAAP Differences Without Country Pair Controls

The main purpose of this paper is to examine whether GAAP differences have any impact on foreign analyst following and forecast accuracy. We begin by estimating the following baseline regressions that control for firm and analyst characteristics but do not control for country characteristics other than GAAP differences in order to utilize the full sample and to ensure that our results hold in regressions without potentially collinear country characteristics included as controls.

$$for\_anal_{ja} = \alpha + \beta_1 gaapdiff_{fa} + \beta_2 mv_j + \beta_3 adr_j + \beta_4 diversity_j + \sum_{i=2}^{13} \eta_i industry_i + \varepsilon_j,$$

where j stands for firm, f for the country location of the firm, and a for the country location of the analyst. The dependent variable is the average annual number of foreign analysts that follow a firm from each of 48 foreign countries during 1998–2004, multiplied by 100 for reporting and expositional purposes. The industry classification we use is the primary industry division in I/B/E/S, which includes a total of 13 sectors. The ADR dummy takes the value of 1 if a firm has ever had an ADR program since 1990, and 0 otherwise. All other variables are the mean of the annual observations for each firm as defined in Table 2.

Table 4 shows the multivariate regression results on analyst following with *gaapdiff1* (column (1)) and *gaapdiff2* (column (2)) as the measure of GAAP differences. Following

---

[20] The industry classification that we use is the primary division in I/B/E/S that includes a total of 13 sectors (including "MISCELL" and "UNDESIG").

[21] As mentioned earlier, we retain all firms with at least one foreign analyst in our full sample because we know of no reason to require more than one foreign analyst. Our main results are qualitatively similar if we require at least five foreign analysts or do not require any foreign analyst to be included in our sample.

TABLE 4
International GAAP Differences and Foreign Analyst Following and Forecast Accuracy

Dependent Variable	*for_anal*	*for_anal*	*afep*	*afep*
gaapdiff1	−0.129 (−3.29)***		−0.079 (−2.55)**	
gaapdiff2		−0.175 (−6.79)***		−0.073 (−2.76)***
mv	0.014 (3.68)***	0.010 (2.97)***	0.015 (5.40)***	0.015 (5.33)***
diversity	0.657 (0.55)	1.859 (1.56)	−31.890 (−1.54)	−31.665 (−1.52)
adr	0.743 (6.83)***	0.905 (8.93)***	−0.251 (−1.98)**	−0.224 (−1.67)*
horizon			−1.935 (−13.78)***	−1.942 (−13.80)***
brsize			0.099 (3.10)***	0.085 (2.44)**
nfirm			0.027 (0.50)	0.026 (0.49)
firmex			0.038 (3.15)***	0.032 (2.63)***
genex			0.024 (2.60)***	0.022 (2.36)**
constant	2.412 (6.52)***	3.111 (9.71)***	−2.060 (−5.15)***	−1.934 (−4.85)***
Industry	Yes	Yes	Yes	Yes
Obs.	291,264	291,264	43,547	43,547
Log Likelihood	−117,239	−116,918		
R^2			0.07	0.08

*, **, *** Significant at 10 percent, 5 percent, and 1 percent, respectively.

The table shows the impact of international GAAP differences on foreign analyst following and foreign analyst forecast accuracy. Foreign analysts are located in a different country from the firms they cover. We identify the geographic location of equity analysts based on *Nelson's Directory of Investment Research* from 1998 to 2004. The main sample excludes firms identified as using either U.S. GAAP or IAS. The dependent variable is for_anal_{ja}, the average annual number of analysts from country *a* who follow firm *j* from another country during 1998–2004. It is 0 if there is no analyst from country *a* following firm *j*. We multiply this variable by 100. The dependent variable $afep_{ijt}$ is absolute forecast error scaled by the most recent stock price in the previous year. To facilitate exposition we multiply this variable by −100 so that higher values represent more accurate forecasts. The annual earnings forecasts are restricted to the latest ones for each firm/year/analyst with a minimum of 30 days before the I/B/E/S report dates during 1998–2004.

All of the variables are defined in Table 2.

We do not report coefficients for industry dummies. Robust z- or t-statistics adjusted for country-pair level clustering are in parentheses.

Rock et al. (2001), we use the negative binomial model to estimate analyst following.[22] In this estimation and all subsequent tests, we adjust standard errors for country-pair level clusters. Cluster analysis inflates standard errors for the potential cross-correlation at the cluster level. Significance levels are much lower with country-pair level cluster adjusted standard errors than without this adjustment (e.g., z-statistics on *gaapdiff1* and *gaapdiff2* are −18.6 and −31.53 without clustering, respectively). This dramatic reduction in reported significance levels should be considered when comparing our results with prior international studies that do not utilize clustered standard errors (e.g., Sarkissian and Schill 2004). The number of analysts following increases with firm size, but is unrelated to forecast diversity. Firms with ADR programs attract more analysts. We include industry dummy variables to control for possible industry effects but do not report the industry coefficients to conserve space.

---

[22] We check the sensitivity of our results by also estimating the model using OLS and logit (both with country-pair clusters), a random effect model, and Tobit regressions. In each case, the coefficient on *gaapdiff1* and *gaapdiff2* is negative and significant at the .01 level. We conclude that our results are not sensitive to the estimation method we use.

For both columns (1) and (2) of Table 4, we find a negative and significant reduction in analyst following as our measures of GAAP differences increase. In interpreting these coefficients, it is important to remember that the coefficient reflects the reduction in the number of foreign analysts from another country per GAAP difference between two countries, but that the negative binomial estimation is nonlinear. We assess the economic magnitude of the effect by looking at the marginal effect of changing the number of GAAP differences while holding other explanatory variables constant. When we set *gaapdiff1* and all of the control variables at their sample mean, the number of foreign analysts *per foreign country* is 0.042, which equals about two foreign analysts in total. When we set *gaapdiff1* at 0 and hold all control variables at the sample mean, the number of foreign analysts *per foreign country* is 0.135, which equals about 6.5 foreign analysts in total, an increase of over 200 percent. For *gaapdiff2* the marginal effect is even larger.

As noted earlier, our regressions include only those firms that have at least one foreign analyst, and we do not include local analysts (located in the same country as the firm) in our analysis. We exclude local analysts because while we think we can identify and control for other factors that might influence analyst behavior *among foreign analysts*, we are less confident that we can identify and control for all factors that might influence local *versus* foreign analyst decisions to follow firms. While GAAP differences might have the most influence in deterring *all* foreign analysts, there are clearly many more issues that could reduce costs for local analysts relative to foreign analysts (e.g., ease of building relationships with management that could result in private information for local analysts). To ensure that our results are not driven by our decision to exclude local analysts, we repeat the analysis including local analysts and setting *gaapdiff1* and *gaapdiff2* to 0 for local analysts. This results in a much larger sample of firms since we include firms that are followed only by local analysts (i.e., we drop the requirement that firms be followed by at least one foreign analyst). Our results are much stronger with local analysts included, with coefficients (z-statistics with country-pair clusters) of −0.325 (−7.82) and −0.287 (−9.80) on *gaapdiff1* and *gaapdiff2*, respectively. We continue reporting results throughout based on the sample that includes only foreign analysts because of our concerns that local analysts have advantages other than GAAP familiarity that we cannot control for in the analysis.

Foreign Analyst Forecast Accuracy and GAAP Differences Without Country Pair Controls

Columns 3 and 4 of Table 4 show our main results for analysts' forecast accuracy, with *afep* (the absolute forecast error scaled by stock price and multiplied by −100 so the variable increases with accuracy and to facilitate reporting and exposition) as the dependent variable.

$$\begin{aligned} afep_{ijt} = {} & \alpha + \beta_1 gaapdiff_{fa} + \beta_2 horizon_{ijt} + \beta_3 \log(brsize_{iY}) + \beta_4 \log(nfirm_{iY}) \\ & + \beta_5 firmex_{ijt} + \beta_6 genex_{it} + \beta_7 mv_{jY} + \beta_8 diversity_{jY} \\ & + \sum_{i=2}^{13} \eta_i industry_i + \sum_{Y=1998}^{2003} \gamma_Y year_Y + \varepsilon_{ijt} \end{aligned}$$

where *i* stands for analyst, *j* stands for firm, *t* for time of the forecast, *f* for country location of the firm, *a* for country location of analyst, *Y* for year. All of the variables are defined in Table 2. We run OLS regressions with country-pair clusters.[23] In contrast to our analysis of analyst following, we utilize our full time-series, cross-sectional data since forecast errors can more appropriately be considered independent over time.

In this analysis, we take the natural log of brokerage size and the number of firms an analyst follows because of skewness in the distribution of these variables.[24] Not surprisingly, the most important determinant of forecast accuracy is forecast horizon. Analysts provide less accurate forecasts when they forecast at longer horizons. The coefficients on research firm size are positive, indicating that better resources available to analysts increase their forecast accuracy. The coefficients on both analysts' firm-specific experience and general forecast experience are positive, suggesting that analysts who are more experienced provide more accurate forecasts. Analysts are more accurate in forecasting larger firms and less accurate in forecasting firms with ADRs. Both GAAP difference measures are negatively related to forecast accuracy, indicating that larger GAAP differences dampen foreign analysts' ability to provide accurate earnings forecasts.[25]

Country-Pair Differences in Other Characteristics and GAAP Differences

Table 4 includes all available observations but does not include controls for country-level characteristics that could potentially impact analyst following and forecast accuracy. Table 5 introduces controls for differences between country-pairs on a number of country-level characteristics that could potentially affect our results, at the cost of modest reductions in sample sizes as not all country characteristics are available for all sample countries. Included in the regressions now are the country-level variables summarized in Table 2. This list includes all the control variables that Sarkissian and Schill (2004) create in order to explain firms' cross-listing decisions as well as the cross-listings between the firm's country and the analysts' country. This comprehensive set of control variables captures differences (or similarities) between the firm's country and the analyst's country in language, colonial history, trade, industrial structure, geographic location, size of the economy and the stock market, stock market liquidity, cross-listing of equity shares, legal origin, and shareholder protection.[26]

---

[23] Once again clustering at the country-pair level has a dramatic effect on significance levels. Without clustering, the t-statistics on *gaapdiff1* and *gaapdiff2* are -15.26 and -16.82, respectively. We also run the regression using firm level rather than country-pair level clustering. Both GAAP difference measures have significantly negative coefficients in this specification.

[24] Our main results remain the same if we use the raw number for brokerage size and number of firms an analyst follows.

[25] In Table 4, the coefficient on *gaapdiff1* is -0.079, which implies that all else equal, when GAAP difference between two countries increases by 1, the absolute forecast error scaled by price increases by 0.079 percent. Since the sample mean absolute forecast error scaled by price is two percent in Table 3, and the mean GAAP difference for *gaapdiff1* is 9, foreign analysts would have been 35 percent $((9*0.079)/2 = 0.35)$ more accurate in their forecasts if there were no GAAP differences between two countries. The coefficient on *gaapdiff2* implies a similar impact on forecast accuracy.

[26] We thank Sergei Sarkissian for sharing the country-pair cross-listing data with us.

We recreate each of their control variables, in most instances using exactly the same measure, since these same measures are expected to explain analysts' coverage decisions and forecast accuracy in addition to our primary experimental variable, GAAP differences.[27] We expect that GAAP differences across countries are also related to a variety of these factors, so failure to control for these variables could lead to a biased estimate of the impact of GAAP differences.[28]

In addition, Table 3 demonstrates that there is substantial variation in the number of analysts we identify as resident in each of our sample countries. This variation reflects differences in the number of analysts who reside in each country and follow foreign firms, and differences in our success in identifying location across countries. Since the number of analysts from each country who follow any given foreign firm is expected to be a function of the total number of analysts in this sample country, we include an additional control variable. We define *tot_for_anal* as the average total number of analysts from the analyst's country who follow at least one foreign firm during the 1998–2004 sample period and include this variable in our analysis of analyst following.[29]

We report the results including these control variables in Table 5. The coefficients on the initial set of control variables are generally consistent with the results reported in Table 4. Despite collinearity among these variables that impacts the significance of individual coefficients, many of the relations between analyst following and these control variables are significant and have intuitively appealing signs. Foreign analyst following is greater when the firm's country and the analyst's country share a common language or colonial history, when economic ties between the countries are stronger, and when the two countries are geographically closer. The negative signs on *gdp* and *capgdp* suggest that analysts' interest flows from richer countries with better-developed stock markets to poorer countries with smaller equity markets. Higher correlation of industrial structure is weakly negatively related to foreign analyst following, contrary to our expectations. There is a strong positive relation between the total number of analysts in the analyst's country and analyst following from that country at the firm level. The most important result in the first two columns of Table 5 is that the relation of GAAP differences to analyst following remains statistically significant across both specifications even after controlling for these additional country characteristics and with country-pair clustering to adjust standard errors. The introduction of the control variables does, however, noticeably decrease the economic significance of the relation between GAAP differences and foreign analyst following to about one-third of the magnitude in our earlier specification. The coefficient on *gaapdiff1*, for example, now implies an 84 percent increase in foreign analyst following as *gaapdiff1* changes from the sample mean of 9 to 0, holding all other variables at the sample mean.

The last two columns of Table 5 show the impact of GAAP differences on foreign analysts' forecast accuracy after we control for country-pair characteristics. In general, the

---

[27] We use a different measure of stock market liquidity that is available for more countries. Sarkissian and Schill (2004) calculate *cap_corr* using ranks rather than actual industry weightings. Our results are similar if we calculate the correlation of industry structure based on ranks.

[28] In an earlier version of the paper we included controls for a variety of country characteristics in levels form, rather than controlling for country-pair differences in these variables. Our measures of GAAP differences were also significantly negatively related to foreign analyst following and forecast accuracy in those earlier tests.

[29] We do not include this variable in our analysis of forecast accuracy since there is no reason to expect a relation between the number of analysts located in the analyst's country and individual analyst forecast accuracy. We did, however, confirm that this relation is insignificant.

TABLE 5
International GAAP Differences and Foreign Analyst Following and Forecast Accuracy, Controlling for Country-Pair Differences

Dependent Variable	*for_anal*	*for_anal*	*afep*	*afep*
gaapdiff1	−0.068		−0.036	
	(−2.02)**		(−1.25)	
gaapdiff2		−0.062		−0.059
		(−2.38)**		(−2.51)**
mv	0.010	0.011	0.014	0.014
	(3.16)***	(3.15)***	(5.38)***	(5.30)***
diversity	−9.127	−9.085	−89.606	−89.403
	(−5.21)***	(−5.23)***	(−2.85)***	(−2.85)***
adr	0.452	0.470	−0.219	−0.189
	(3.38)***	(3.66)***	(−1.74)*	(−1.48)
horizon			−1.878	−1.876
			(−13.98)***	(−13.96)***
brsize			0.140	0.132
			(5.25)***	(4.91)***
nfirm			0.020	0.023
			(0.48)	(0.58)
firmex			0.044	0.041
			(3.43)***	(3.37)***
genex			0.020	0.019
			(1.93)*	(1.80)*
dist	−0.155	−0.162	−0.019	−0.021
	(−6.25)***	(−6.32)***	(−0.56)	(−0.63)
cult	1.331	1.315	0.313	0.253
	(5.19)***	(5.27)***	(1.16)	(1.11)
econ_prox	0.172	0.162	0.007	0.006
	(2.69)***	(2.61)***	(1.10)	(0.86)
cap_corr	−0.604	−0.609	0.442	0.428
	(−1.85)*	(−1.85)*	(1.24)	(1.31)
stkturn	−0.029	−0.015	0.119	0.119
	(−0.49)	(−0.24)	(1.07)	(1.09)
gdp	−0.007	−0.006	0.006	0.006
	(−1.98)**	(−1.64)	(1.07)	(1.23)
capgdp	−0.016	−0.015	0.018	0.025
	(−3.13)***	(−3.04)***	(1.59)	(2.05)**
legorig	−0.194	−0.187	0.066	0.070
	(−0.87)	(−0.84)	(0.31)	(0.32)
antidir	−0.062	−0.046	0.087	0.084
	(−0.88)	(−0.68)	(1.15)	(1.07)
crosslist	20.039	16.498	−0.276	−0.582
	(0.76)	(0.74)	(−0.30)	(−0.63)
tot_for_anal	0.005	0.005		
	(4.92)***	(5.12)***		
Constant	1.638	1.714	−3.144	−2.836
	(3.41)***	(3.22)***	(−4.56)***	(−4.68)***

(*continued on next page*)

TABLE 5 (continued)

Dependent Variable	*for_anal*	*for_anal*	*afep*	*afep*
Industry	Yes	Yes	Yes	Yes
Obs.	212,121	212,121	41,239	41,239
Log Likelihood	−103,927	−103,922		
R^2			0.09	0.10

*, **, *** Significant at 10 percent, 5 percent, and 1 percent, respectively.
The table shows the impact of international GAAP differences on foreign analyst following and foreign analyst forecast accuracy. Foreign analysts are located in a different country from the firms they cover. We identify the geographic location of equity analysts based on *Nelson's Directory of Investment Research* from 1998 to 2004. The main sample excludes firms identified as using either U.S. GAAP or IAS. The dependent variable is for_anal_{ja}, the average annual number of analysts from country a who follow firm j from another country during 1998–2004. It is 0 if there is no analyst from country a following firm j. We multiply this variable by 100. The dependent variable $afep_{ijt}$ is absolute forecast error scaled by the most recent stock price in the previous year. To facilitate exposition we multiply this variable by -100 so that higher values represent more accurate forecasts. The annual earnings forecasts are restricted to the latest ones for each firm/year/analyst with a minimum of 30 days before the I/B/E/S report dates during 1998–2004.
All of the variables are defined in Table 2.
We do not report coefficients for industry dummies. Robust z- or t-statistics adjusted for country-pair level clustering are in parentheses.

control variables do not have significant incremental explanatory power for forecast accuracy. This lack of significance could indicate that while these controls affect analyst coverage decisions, once analysts decide to cover a firm they overcome the effects of language, distance, etc., in order to provide accurate forecasts, but should be interpreted cautiously due to collinearity among the control variables that could dampen significance levels. The addition of these controls reduces the significance of the negative coefficient on *gaapdiff1* below conventional levels, while *gaapdiff2* remains significantly negatively related to forecast accuracy.[30] This attenuation in significance level may reflect the fact that analysts' incentives to provide accurate forecasts motivates them to incur costs to overcome the effects of GAAP differences once they have decided to follow a foreign firm, moderating the impact of GAAP differences on forecast accuracy. In part, the decision to follow a foreign firm signals the analyst's decision that the benefits of following the firm outweigh the costs, including the costs of dealing with unfamiliar GAAP.[31]

---

[30] The economic significance of the effect of our GAAP difference variables on forecast accuracy is only about half as big after controlling for country characteristics.

[31] We also investigate the possibility that analysts facing greater GAAP differences minimize their forecast errors by following analysts facing fewer GAAP differences by calculating the lead-follow ratio (LFR) developed by Cooper et al. (2001). Put simply, the LFR is the ratio of the total number of days between an analyst's forecast and the two preceding forecasts to the total number of days between an analyst's forecast and the two subsequent forecasts. Leading analysts are expected to have their forecasts quickly followed by others, so leaders are expected to have ratios greater than 1. Followers have lower LFRs than leaders. If analysts facing greater GAAP differences are followers there should be a negative correlation between our GAAP difference measures and the LFR (higher GAAP differences should be related to followers with lower LFR.) Both of our GAAP difference measures are negatively correlated with the LFR at around the −0.05 level (−0.0492 with *gaapdiff1* and −0.05661 for *gaapdiff2*). While this is consistent with the idea that analysts facing greater GAAP differences follow analysts facing fewer GAAP differences, regressions using all country-pair characteristics as controls do not reveal a significant negative relation between either GAAP difference measure and the LFR.

Alternative Measures of Accounting Differences and Firm-Level Accounting Measures

Adding CIFAR and Earnings Management Differences

One criticism of our approach to measuring accounting differences across countries is that we focus on the accounting standards in place in each country rather than focusing on differences in actual accounting practice across countries. Since the primary issue in accounting harmonization debates deals with whether to harmonize accounting standards, we believe that this focus is appropriate. In addition, since understanding the accounting rules in place seems like a prerequisite to effectively following and forecasting a firm, we believe that variation in rules likely has a significant effect on analysts, even if there is additional variation across or within countries in how GAAP are implemented. Nevertheless, we also estimated our equations that include all the country controls mentioned above and included two alternative measures of accounting variation across countries, CIFAR disclosure scores, which are based on disclosure practices in place at large companies in each country, and the earnings management ranks constructed by Leuz et al. (2003). In both cases, we calculate the difference between the firm's country and the analyst's country to capture how different the disclosure or earnings management practices are between each country-pair.

Untabulated results show that the CIFAR disclosure index has no incremental explanatory power over our measure of GAAP differences, and the inclusion of the CIFAR index does not alter our inferences with respect to *gaapdiff1* or *gaapdiff2*. Earnings management, on the other hand, has a significantly negative relation with analyst following. Since the earnings management rank increases with greater earnings management, this negative relation indicates that analyst following falls when there is greater earnings management in the firm's country than in the analyst's country, as intuition would suggest. While including this variable suggests that variation in accounting practices does impact foreign analysts, both *gaapdiff1* and *gaapdiff2* remain significantly negatively related to foreign analyst following, suggesting that variation in both accounting standards and practices across countries affect analysts. We also perform similar analyses on the forecast accuracy sample. Neither CIFAR nor earnings management is significantly related to forecast accuracy. None of our earlier inferences are changed except *gaapdiff2* loses significance when earnings management is added to the equation.

Adding Firm-Level Controls for Auditor and Audit Opinion

Our *gaapdiff* variables measure GAAP differences at the country level but fail to capture variation in how those standards are implemented within the country at the firm level. Testing whether firm-level variation in the implementation of accounting standards also affects analysts requires a measure of standards implementation at the firm level, and such measures are not widely available. We develop two proxies for firm-level implementation of accounting standards that are widely available. We create a variable called *Big4* that takes on a value of 1 if the firm's auditor is one of the Big 4 accounting firms, and 0 otherwise.[32] We also create a variable called *qualified* that takes a value of 1 if the firm receives a qualified audit opinion in Year *Y*. We include each of these variables in our regressions that include country-pair control variables. If better firm-level implementation of accounting standards attracts analysts and makes it easier to forecast earnings, then we expect a positive coefficient on *Big4* and a negative coefficient on *qualified*. Auditor identity and audit opinion are not available for all sample firms so our samples for these tests are smaller than the full sample.

---

[32] Auditor identity is a static variable in the Worldscope database so we are not able to classify each firm-year based on auditor identity.

Untabulated results show that while *Big4* has the expected positive sign and *qualified* has the expected negative sign, these variables are not significantly related to analyst following. Including these variables does not alter our inferences about country-level GAAP differences since both measures of GAAP differences are significantly negatively related to foreign analyst following in these tests. In regressions explaining forecast accuracy *Big4* has the expected positive sign and is significant and *qualified* has the expected negative sign and is also significant. The variable *gaapdiff1* remains insignificantly related to forecast accuracy, while the significant negative relation between *gaapdiff2* and forecast accuracy continues to hold in these tests. In sum, adding proxies for variation in the implementation of accounting standards at the firm level does not alter our inferences.

Sensitivity Tests

In this section, we report the results of a number of additional tests that examine the sensitivity of our results to alternative specifications. Because our main results for forecast accuracy are not as strong as the results for analyst following, the focus of this analysis is on alternative specifications for our tests of analyst following.

Alternative Proxies for Dependent Variables

So far we have used one proxy for foreign analyst following, *for_anal*, and one proxy for forecast accuracy, *afep*. We also estimated our equations with alternative proxies for our dependent variables. Our alternative proxy is *for_anal/total*, which scales the number of analysts as previously defined by the average total number of foreign analysts following each firm during 1998–2004. The scaling effectively converts this alternative proxy for analyst following to the proportion of total analysts following the firm who are from each foreign country. An advantage of this specification is that it effectively controls for firm characteristics that impact analysts' willingness to follow the firm in general, without regard to analyst location. Because this scaling method controls for firm characteristics, we do not include any firm characteristics in this specification. The results for analyst following are reported in columns (1) and (2) of Table 6. Both GAAP difference measures remain significant and negatively related to analyst following in this specification.

Our main results for analyst following use firm-level data. One concern with this specification is that there could be cross-sectional correlation between error terms at the country-pair level. Our main results include clustered standard errors at the country-pair level to mitigate this concern, but an alternative way to deal with the problem is to collapse the data down to country-pair observations based on the mean value for the firms in each country. This specification not only eliminates concerns about repeated observations at the country-pair level, but also only utilizes one observation for each country-pair and addresses concerns that our results could be driven by the concentration of observations in a few key countries. Due to missing values for some country-pair control variables, our sample size for this analysis is 1,406 country-pair observations. In this analysis, the dependent variable and all firm-level control variables reflect the mean across all firms located in each country. Because the alternative proxy for analyst following discussed above also controls for firm characteristics, we perform this analysis using the alternative proxy.

We report this country-pair analysis in columns (3) and (4) of Table 6. As shown in this table, both GAAP difference measures are negative and significant in this specification. We also repeated this country-pair analysis using our original proxy for the number of

TABLE 6
International GAAP Differences and Foreign Analyst Following, Alternative Specifications

	Alternative Proxy for Analyst Following		Country-Pair Level Alternative Proxy		Including Firms with No Foreign Analyst		Dropping U.S., U.K., and Canada		Developed Countries Only		Matched Sample	
Model	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
gaapdiff1	−0.075 (−1.95)*		−0.061 (−3.01)***		−0.084 (−2.09)**		−0.080 (−1.99)**		−0.130 (−3.76)***		−0.035 (−4.12)***	
gaapdiff2		−0.071 (−2.25)**		−0.109 (−6.10)***		−0.061 (−1.87)*		−0.101 (−2.48)**		−0.080 (−3.01)***		−0.027 (−3.22)***
mv					0.068 (6.06)***	0.069 (6.40)***	0.003 (0.82)	0.005 (1.08)	0.020 (4.68)***	0.022 (4.46)***	0.006 (3.89)***	0.004 (3.48)***
diversity					0.055 (0.88)	0.047 (0.76)	7.348 (0.39)	5.998 (0.32)	−7.081 (−4.37)***	−7.026 (−4.31)***	−166.701 (−0.75)	14.772 (0.10)
adr					1.003 (7.38)***	1.02 (7.80)***	0.716 (5.83)***	0.738 (5.70)***	0.683 (4.73)***	0.696 (5.60)***	0.212 (3.65)***	0.188 (4.37)***
dist	−0.112 (−3.93)***	−0.120 (−4.15)***	−0.191 (−11.40)***	−0.202 (−11.87)***	−0.144 (−5.23)***	−0.151 (−5.61)***	−0.290 (−8.89)***	−0.294 (−9.40)***	−0.111 (−3.64)***	−0.114 (−3.78)***	−0.000 (−5.81)***	−0.000 (−7.22)***
cult	1.423 (4.74)***	1.383 (4.80)***	1.379 (8.83)***	1.301 (8.61)***	1.133 (3.77)***	1.173 (3.98)***	1.291 (4.78)***	1.383 (5.16)***	0.964 (3.34)***	1.135 (3.52)***	−0.132 (1.22)	0.046 (0.46)
econ_prox	0.198 (2.61)***	0.184 (2.52)**	0.106 (5.71)***	0.100 (5.57)***	0.125 (2.23)**	0.117 (2.15)**	0.173 (3.33)***	0.152 (3.31)***	0.175 (2.48)**	0.163 (2.27)**	0.009 (2.74)***	0.005 (1.81)*
cap_corr	−0.720 (−2.05)**	−0.714 (−2.03)**	0.224 (1.19)	0.184 (0.98)	−0.566 (−1.74)*	−0.561 (−1.73)*	0.074 (0.23)	0.108 (0.33)	−0.109 (−0.35)	−0.287 (−0.89)	0.009 (0.08)	0.208 (1.59)
stkturn	−0.048 (−0.82)	−0.031 (−0.52)	0.115 (−2.01)**	−0.113 (−1.90)*	−0.004 (−0.05)	0.007 (0.11)	−0.240 (−2.62)***	−0.262 (−2.91)***	0.556 (2.33)**	0.482 (1.93)*	0.036 (0.83)	0.039 (0.87)
gdp	−0.007 (−1.91)*	−0.006 (−1.54)	0.002 (0.38)	0.002 (0.39)	−0.011 (−3.63)**	−0.011 (−3.35)***	−0.064 (−4.90)***	−0.064 (−4.91)***	−0.017 (−4.57)***	−0.015 (−4.04)***	0.006 (2.28)**	0.010 (3.67)***
capgdp	−0.019 (−3.21)***	−0.018 (−3.12)***	−0.014 (−1.96)**	−0.011 (−1.58)	−0.013 (−2.65)***	−0.013 (−2.61)***	−0.012 (−2.55)**	−0.012 (−2.62)***	−0.008 (−1.59)	−0.006 (−1.23)	0.002 (0.37)	0.003 (0.72)

(*continued on next page*)

TABLE 6 (continued)

legorig	−0.116	−0.105	−2.698	−3.698	−0.224	−0.202	13.413	13.281	−4.678	−0.110	0.306	−0.202
	(−0.45)	(−0.41)	(−0.55)	(−0.82)	(−0.91)	(−0.82)	(0.58)	(0.67)	(−1.47)	(−0.01)	(0.93)	(−0.85)
antidir	0.010	0.033	−0.010	−0.066	−0.084	−0.076	−0.682	−0.653	−0.854	−0.897	0.148	0.107
	(0.13)	(0.43)	(−0.07)	(−0.50)	(−1.04)	(−0.94)	(−2.54)**	(−2.38)**	(−3.11)***	(−3.27)***	(1.27)	(1.08)
crosslist	21.971	15.162	−0.045	−0.056	32.871	28.48	−0.030	0.001	0.074	0.107	−0.029	−0.034
	(0.71)	(0.58)	(−1.04)	(−1.28)	(1.26)	(1.29)	(0.34)	(0.01)	(0.96)	(1.39)	(−1.08)	(−1.48)
tot_for_anal	0.005	0.005	0.006	0.006	0.004	0.004	0.025	0.024	0.004	0.004	0.001	0.001
	(4.48)***	(4.84)***	(11.47)***	(11.79)***	(4.28)***	(4.36)***	(5.26)***	(5.16)***	(4.94)***	(5.02)***	(3.13)***	(4.32)***
Constant	0.308	0.438	0.596	1.392	0.943	0.863	2.281	2.834	1.049	0.802	1.168	1.001
	(0.60)	(0.74)	(2.16)**	(4.67)***	(1.90)*	(1.63)	(3.94)***	(4.18)***	(2.10)**	(1.30)	(4.90)***	(3.00)***
Industry	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs	212,121	212,121	1,406	1,406	693,861	693,861	65,892	65,892	105,160	105,160	14,126	15,113
Log Likelihood	−92,106	−92,096	−1,470	−1,455	−127,473	−127,484	−33,035	−33,017	−75,762	−75,801	−33,379	−31,405

---

*, **, *** Significant at 10 percent; 5 percent; and 1 percent, respectively.

The table shows the impact of international GAAP differences on foreign analyst following. Foreign analysts are located in a different country from the firms they cover. We identify the geographic location of equity analysts based on *Nelson's Directory of Investment Research* from 1998 to 2004. The main sample excludes firms identified as using either U.S. GAAP or IAS. For columns (5)–(12), the dependent variable is for_anal_{ja}, the average annual number of analysts from country *a* who follow firm *j* from another country during 1998–2004. It is 0 if there is no analyst from country *a* following firm *j*. We multiply this variable by 100.

All of the variables are defined in Table 2. We do not report coefficients for industry dummies.

Robust z-statistics adjusted for country-pair level clustering (this adjustment does not apply to columns (3) and (4)) are in parentheses. Columns 1 and (2) use an alternative measure of analyst following that scales analyst following by the total number of foreign analyst following the firm, columns (3) and (4) use the alternative proxy for analyst following at the country-pair level rather than firm level, columns (5) and (6) report results for the sample that include firms with 0 foreign analyst, columns (7) and (8) report results for the sample excluding the U.S., the U.K. and Canada, columns (9) and (10) report results using observations from 24 developed countries, columns (11) and (12) report the results on matched samples (*gaapdiff1 and gaapdiff2* are binary variables in the matched samples).

analysts following where we do not scale the variable by the total number of foreign analysts following the firm. We include firm characteristics in this regression. Again, both GAAP difference variables retained negative and significant signs.

The alternative measure of forecast accuracy we use is *afea*, the absolute forecast error scaled by actual earnings as reported in I/B/E/S. We also calculated forecast accuracy by subtracting the mean or the median forecast 30 days before the I/B/E/S report date, rather than actual earnings, from each forecast. Neither modification qualitatively alters our main results, except both *gaapdiff1* and *gaapdiff2* are negative and significantly related to *afea*, while only *gaapdiff2* is significantly related to our original proxy, *afep* (these results are not tabulated).

Including Firms with No Foreign Analyst Following

Our primary results use a sample of firms with at least one foreign analyst following the firm. Since our primary interest is in explaining variation in analyst following across foreign countries, this sample seems appropriate. To examine the sensitivity of our results to this sample restriction, we also estimated our analyst following regressions on a sample including all firms, regardless of whether the firm is followed by any foreign analyst. Since many firms do not attract foreign analysts, this sample is much larger than our primary sample. We report the results of this analysis in columns (5) and (6) of Table 6. Using this expanded sample does not qualitatively alter the conclusions from our analysis.

Dropping the U.S., the U.K., and Canada, Retaining Only Developed Countries

One potential concern is that differences between more developed and less developed countries are responsible for our results. While our model includes a comprehensive set of country-pair control variables that should control for important differences between more developed and less developed countries, it is still possible that our results reflect the fact that analysts and the firms they follow tend to be located in the more developed countries, which tend to have similar GAAP. If so, then analyst following could be concentrated in the more developed countries and our GAAP difference measures could simply proxy for other omitted factors that distinguish more from less developed countries. We perform two additional sets of analyses to address this concern. First, the U.K., the U.S., and Canada are all highly developed countries with significant concentrations of both analysts and firms in our sample. These three countries share a number of historic, political, and economic ties, and all have similar GAAP. It is possible that the concentration of sample observations in these three countries is responsible for our results. Accordingly, we exclude these three countries from our sample and re-estimate our model of analyst following including our controls for country-pair characteristics. The results of this specification are reported in columns (7) and (8) of Table 6. Not surprisingly, deleting these three countries has a substantial impact on our sample size, causing a loss of around two-thirds of our sample observations. Nevertheless, both GAAP difference measures retain negative and significant coefficients in this specification, leading us to conclude that the inclusion of these three key countries in our sample is not responsible for our results.

We also reran our analyst following regressions including *only* the firm and analyst observations from the more developed countries.[33] Columns 9 and 10 of Table 6 report the results of this analysis. Again, in tests that control for country-pair characteristics, both

---

[33] The 24 developed (out of our total of 49) countries we retained in this analysis are: Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Hong Kong, Ireland, Italy, Japan, Luxembourg, The Netherlands, New Zealand, Norway, Portugal, Singapore, Spain, Sweden, Switzerland, the U.K., and the U.S.

gaapdiff1 and *gaapdiff2* have significantly negative coefficients. We conclude that our results are not due to GAAP differences simply capturing variation in analyst location across more developed versus less developed countries.

Utilizing a Matched Sample for Firms Followed by Two or More Analysts

We also construct a matched sample in which each firm in each year is followed by analysts from at least two countries with different *gaapdiff1* (or *gaapdiff2*). For each firm and year, we partition the sample into a high GAAP difference group and a low GAAP difference group based on the median score of the GAAP difference measure for that firm and year. One potential advantage of using such a matched sample is that it controls for firm effects that could impact analysts. Another advantage of this approach is that the resulting sample contains no zero analyst following observations. This specification therefore tests whether our measures of GAAP differences explain the number of analysts following the firm rather than just explaining whether any analyst from any given country follows the firm. There are 1,524 distinct firms for this matched sample. Both mean and median tests show that there are significantly more foreign analysts from countries with lower GAAP difference. For example, the average number of analysts following from low GAAP difference countries using *gaapdiff1* and *gaapdiff2*, respectively, are 5.58 and 5.98, while the average number of analyst following from high GAAP difference countries are only 2.72 and 2.07, respectively. The median number of analyst following from low GAAP difference countries using *gaapdiff1* and *gaapdiff2*, respectively, are four and five, while the median number of analyst following from high GAAP difference countries are only one and two, respectively. The differences are significant at the 1 percent level. The forecasts of foreign analysts from low GAAP difference countries are also more accurate than those of foreign analysts from high GAAP difference countries, and these results are significant at conventional levels with both measures.

We also repeat our regression analysis on this sample of firms to control for our country-pair control variables. The actual number of GAAP differences between the firm's home country and the analyst's home country varies across firm-country-pairs in both the high and low GAAP difference groups because these partitions are based on firm-year median GAAP differences. Accordingly, in these regressions, the *gaapdiff* measures take on only two values, 0 for the low GAAP difference group and 1 for the high GAAP difference group. We then run regressions that control for all of previously identified country characteristics and include the binary variable, *gaapdiff1* or *gaapdiff2*. Columns 11 and 12 of Table 6 report the results of running our analyst following regressions on the matched sample. Both of our GAAP difference measures retain negative and significant coefficients in this matched sample. We conclude that neither uncontrolled firm characteristics nor the presence of zero observations is affecting our inferences in the paper.

Time-Series Analysis on Firms Voluntarily Adopting International Accounting Standards

Despite extensive sensitivity analyses on our cross-sectional sample, it is of course always possible that there is some omitted variable that we have not identified that our GAAP difference measures could proxy for. In addition, our cross-sectional analysis is subject to some concerns due to the potential endogeneity of accounting standards at the national level. Accordingly, we also provide analysis of a sample of firms switching from local standards to IAS. While this analysis has some advantages, it is not clear that the time-series approach is superior to the cross-sectional approach utilized so far. There are several considerations. First, an examination of firms voluntarily switching to IAS necessarily involves a much smaller sample. Second, firms switching to IAS are a self-selected sample of firms that are likely to be accessing capital markets and undertaking a wide range of activities to attract analyst and investor interest in addition to changing accounting standards, so this approach introduces a new set of potentially omitted variables. Third, it is not clear exactly how to code and interpret "GAAP differences" for firms switching to IAS. For local analysts not familiar with IAS, this switch may increase GAAP differences. For foreign analysts not familiar with IAS, this switch could increase or decrease GAAP differences, depending on whether the firm's local GAAP or IAS is closer to the foreign analyst's home country GAAP. For analysts familiar with IAS, the switch is likely to reduce or eliminate any GAAP induced costs of following the firm. Accordingly, our analysis of this sample focuses on the effects of switching to IAS on two groups of analysts—those following other firms using IAS and those not following other firms using IAS. Unlike our earlier tests that focus only on foreign analysts, we include both local and foreign analysts in this time-series analysis.

We construct a sample of firms switching from their local GAAP to IAS based on accounting standards coding contained in Worldscope for 1998–2004. Because of known errors in coding accounting standards use, we eliminate all firms that we consider "dubious" IAS adopters.[34]

Firms switching to IAS between 1998 to 2004	1,096
Less: Firms missing analyst following data in year −1 or zero	(652)
Less: "Dubious" IAS adopters	
Firm adopts IAS more than once (after reverting to local GAAP)	(41)
Firm switches back to local standards permanently after switching to IAS	(24)
Firm first switches to U.S. GAAP and then to IAS	(6)
Final Sample	373

We then examine changes in analyst following for these 373 firms between the year before IFRS adoption (−1) and the year of IAS adoption. For purposes of this analysis, we put analysts into two groups, those following a firm using IAS at year −1, and those not following a firm using IAS at year −1. If GAAP differences matter to analysts, then the prediction is that analyst following will increase among analysts with IAS experience, and stay constant or decline among the analysts without IAS experience. That is exactly what we find. Between year −1 and zero, the mean (median) analyst following for analysts with IAS experience goes from 4.49 (2) to 5.35 (3). This change in analyst following is significant at the 5 percent level for the mean, 1 percent level for the median. For analysts without IAS experience, the mean (median) number of analysts following goes from 5.32 (3) to 5.11 (3) between years −1 and 0. Neither of these changes is significant. These results corroborate our cross-sectional analyses and suggest that adopting IAS reduces costs for analysts who are familiar with IAS, but not for other analysts. These tests also help allay concerns that the potential endogeneity of accounting standards at the country level is impacting our inferences.

VI. CONCLUSIONS

This paper investigates whether GAAP differences across countries are associated with foreign analyst following and foreign analysts' forecast accuracy. We compute two proxies for GAAP differences for 1,176 country-pairs and provide evidence that GAAP differences are negatively associated with both foreign analyst following and forecast accuracy. Our results are quite robust across alternative model specifications and sample selections for analyst following, though the results for forecast accuracy are more sensitive to modifications in our research design and sample composition. Our results hold when we control for a wide range of country-pair characteristics in addition to GAAP differences. Despite our attempts to control for the other country-pair characteristics, we caution that we are unable to completely rule out the possibility that our GAAP difference measures capture other country-pair differences. However, supplemental time-series analysis on a sample of firms voluntarily adopting IAS yields similar conclusions, enhancing our confidence that GAAP differences are responsible for our results.

---

[34] For example, see Covrig et al. (2007, footnote 21).

The evidence that GAAP differences matter even to financial analysts who specialize in collecting, analyzing, and disseminating financial information about the covered firms suggests that there are potential economic costs associated with variation in accounting rules across countries. We again caution, however, that our research by no means provides a comprehensive assessment of the costs or benefits of accounting harmonization. Our research does point out the need for additional study of the costs and benefits of accounting harmonization to provide important policy inputs to standard setting bodies around the world. The recent adoption of IAS throughout the European Union provides an opportunity to extend our results to a setting where time-series analysis around a mandated change in accounting could be undertaken.

APPENDIX
ACCOUNTING STANDARDS INCLUDED IN *gaapdiff2*

The IAS numbers used for reference in this table are the primary IAS standard reference given in the GAAP 2001 Survey used for coding each item. Since the construction of *gaapdiff2* is more mechanical and less judgmental than *gaapdiff1*, we ignored the textual descriptions for each item and simply coded each country with a 1 if the IAS section referred to is mentioned for that country. Since we are interested in GAAP differences across countries, *gaapdiff2* is the sum of the number of differences observed in the scores across all 52 items for each country-pair.

IAS Standard No.	Area of Accounting
IAS No. 1.7	Statement of changes in equity
IAS No. 7	Cash flow statement
IAS No. 7.45	Reconciliation of cash flow to cash and cash equivalents
IAS No. 8.6	Extraordinary items
IAS No. 10	Subsequent events
IAS No. 11	Construction contracts
IAS No. 11.22	Contract method for construction contracts
IAS No. 12	Deferred taxes
IAS No. 12.53	Discounting of deferred taxes
IAS No. 14	Segment reporting
IAS No. 14.26	Basis of determining segments
IAS No. 14.55	Segment reporting of assets and liabilities
IAS No. 16	Asset revaluation
IAS No. 17	Leases
IAS No. 17.12	Finance leases
IAS No. 17.25	Operating lease expenses

IAS No. 17.30	Finance lease income
IAS No. 19	Employee benefit obligations
IAS No. 19.52	Constructive obligation under pension plans
IAS No. 19.64	Pension obligation measurement
IAS No. 19.92	Actuarial gains and losses on Pensions
IAS No. 2.36	Asset impairment testing
IAS No. 21.11	Foreign currency monetary balances
IAS No. 21.36	Price level adjustment in hyperinflationary economies
IAS No. 21.37	Disposal of foreign segments
IAS No. 22	Classification of business combinations on acquisition
IAS No. 22.31	Provisions in the context of acquisitions
IAS No. 22.8	Business combinations
IAS No. 24	Related-party transactions
IAS No. 28	Accounting for associates
IAS No. 29	Inflation accounting
IAS No. 31	Joint ventures
IAS No. 31.25	Joint ventures < 20 percent owned
IAS No. 32.18	Classification of issuer's financial instruments
IAS No. 32.23	Compound financial instruments
IAS No. 32.77	Disclosure of fair values of financial assets and liabilities
IAS No. 33	Earnings per share
IAS No. 35	Discontinued operations
IAS No. 36	Asset impairment
IAS No. 36.5	Conditions to record asset impairment
IAS No. 37	Provisions
IAS No. 37.45	Discounting of provisions
IAS No. 38	Intangible assets
IAS No. 38.42	Research and development
IAS No. 38.51	Capitalization of internally generated intangibles such as brands
IAS No. 38.56	Start-up costs
IAS No. 38.57	Training advertising and pre-opening costs
IAS No. 38.99	Annual impairment reviews on long-lived assets
IAS No. 39.142	Hedge accounting for derivatives
IAS No. 39.35	Derecognition of financial assets
IAS No. 39.93	Trading liabilities and derivatives
IAS No. 40.69	Fair values of investment properties

REFERENCES

Ang, J. S., and S. J. Ciccone, 2001. International differences in analyst forecast properties. Working paper, University of New Hampshire.

Ashbaugh, H., and M. Pincus. 2001. Domestic accounting standards, international accounting standards, and the predictability of earnings. *Journal of Accounting Research* 39: 417–434.

Bae, K-H., R. M. Stulz, and H. Tan. 2007. Do local analysts know more? A cross-country study of the performance of local analysts and foreign analysts. *Journal of Financial Economics* (forthcoming).

Ball, R., A. Robin, and J. S. Wu. 2003. Incentives versus standards: Properties of accounting income in four East Asian countries. *Journal of Accounting and Economics* 36 (1–3): 235–270.

———. 2006. International financial reporting standards (IFRS): Pros and cons for investors. *Accounting and Business Research* (June): 5–27.

Bank of New York. 2001. *ADR Directory*. Available at: http://160.254.123.37/dr_directory.jsp.

Barniv, R., M. Myring, and W. Thomas. 2005. The association between the legal and financial reporting environments and forecast performance of individual analysts. *Contemporary Accounting Research* 4 (Winter): 727–758.

Barth, M., G. Clinch, and T. Shibano. 1999. International accounting harmonization and global equity markets. *Journal of Accounting and Economics* 26: 201–235.

Basu, S., L. Hwang, and C-L. Jan. 1998. International variation in accounting measurement rules and analysts' earnings forecast errors. *Journal of Business Finance & Accounting* 25: 1207–1247.

Bekaert, G., C. R. Harvey, and C. Lundblad. 2005. Does financial liberalization spur growth? *Journal of Financial Economics* 77: 3–55.

Bradshaw, M., B. Bushee, and G. Miller. 2004. Accounting choice, home bias, and U.S. investment in non-U.S. firms. *Journal of Accounting Research* 42: 795–841.

Bushman, R. M., J. D. Piotroski, and A. J. Smith. 2004. What determines corporate transparency? *Journal of Accounting Research* 42: 207–252.

———, ———, and ———. 2005. Insider trading restrictions and analysts' incentives to follow firms. *Journal of Finance* 60: 35–66.

Chang, J. J., T. Khanna, and K. Palepu. 2000. *Analyst Activity around the World*. HBS Strategy Unit Working Paper No. 01-061 (January). Available at: http://ssrn.com/abstract=204570.

Clement, M. B. 1999. Analyst forecast accuracy: Do ability, resources, and portfolio complexity matter? *Journal of Accounting and Economics* 27: 285–303.

Comprix, J., K. A. Muller III, and M. Stanford. 2003. Economic consequences from mandatory adoption of IASB standards in the European Union. Working paper, The Pennsylvania State University.

Cooper, R., T. Day, and C. Lewis. 2001. Following the leader: A study of individual analysts' earnings forecasts. *Journal of Financial Economics* 61: 383–416.

Covrig, V., M. L. DeFond, and M. Hung. 2007. Home bias, foreign mutual fund holdings, and the voluntary adoption of international accounting standards. *Journal of Accounting Research* 25: 41–70.

Ding, Y., O-K. Hope, T. Jeanjean, and H. Stolowy. 2007. Differences between domestic accounting standards and IAS: Measurement, determinants and implications. *Journal of Accounting and Public Policy* 26 (1): 1–38.

Guan, Y., O-K. Hope, and T. Kang. 2006. Does similarity of local GAAP to U.S. GAAP explain analysts' forecast accuracy? *Journal of Contemporary Accounting and Economics* 2 (2) (December).

Hope, O-K. 2003a. Accounting policy disclosures and analysts' forecasts. *Contemporary Accounting Research* 20 (Summer): 295–321.

———. 2003b. Disclosure practices, enforcement of accounting standards, and analysts' forecast accuracy: An international study. *Journal of Accounting Research* 41: 235–272.

Hung, M. 2001. Accounting standards and value relevance of financial statements: An international analysis. *Journal of Accounting and Economics* 30: 401–420.

Kini, O., S. L. Mian, M. Rebello, and A. Venkateswaran. 2005. *Core Specialization, Forecast Accuracy, and the Market for International Analysts*. Available at: http://ssrn.com/abstract=849306.

Lang, M. H., K. V. Lins, and D. P. Miller. 2003. ADRs, analysts, and accuracy: Does cross-listing in the United States improve a firm's information environment and increase market value? *Journal of Accounting Research* 41 (2): 317–345.

———, ———, and ———. 2004. Concentrated control, analyst following, and valuation: Do analyst matter most when investors are protected least? *Journal of Accounting Research* 42: 589–623.

La Porta, R., F. Lopez-de-Silanes, A. Shleifer, and R. W. Vishny. 1998. Law and finance. *Journal of Political Economy* 106 (6): 1113–1155.

Leuz, C., D. Nanda, and P. D. Wysocki. 2003. Investor protection and earnings management: An international comparison. *Journal of Financial Economics* 69: 505–527.

Nobes, C. 2001. GAAP 2001—A survey of national accounting rules benchmarked against international accounting standards. *International Forum on Accountancy Development* (IFAD).

Rock, S., S. Sedo, and M. Willenborg. 2001. Analyst following and count-data econometrics. *Journal of Accounting and Economics* 30: 351–373.

Sarkissian, S., and M. J. Schill. 2004. The overseas listing decision: New evidence of proximity preference. *Review of Financial Studies* 17 (3): 769–809.

Young, D., and D. A. Guenther. 2003. Financial reporting environments and international capital mobility. *Journal of Accounting Research* 41: 553–579.

THE ACCOUNTING REVIEW
Vol. 83, No. 3
2008
pp. 629–664

Fair Value Accounting for Liabilities and Own Credit Risk

Mary E. Barth
Stanford University

Leslie D. Hodder
Indiana University

Stephen R. Stubben
The University of North Carolina at Chapel Hill

ABSTRACT: We find that equity returns associated with credit risk changes are attenuated by the debt value effect of the credit risk changes, as Merton (1974) predicts. We find that the relation between credit risk changes and equity returns is significantly less negative for firms with more debt—controlling for asset value changes, credit risk increases (decreases) are associated with equity value increases (decreases). This result obtains across credit risk levels. The relation is associated with changes in both expected cash flows and systematic risk, as reflected in analyst earnings forecasts and equity cost of capital. By inverting the Merton (1974) model, we provide descriptive evidence that if unrecognized debt value changes were recognized in income, but not unrecognized asset value changes, most credit upgrade (downgrade) firms would recognize lower (higher) income. These potentially counterintuitive income effects primarily are attributable to incomplete recognition of contemporaneous asset value changes. However, for a substantial majority of downgrade firms we find that recognized asset write-downs exceed unrecognized gains from debt value decreases. This mitigates concerns that income effects from recognizing changes in debt values would be anomalous for such firms.

Keywords: *fair value accounting; income recognition; debt; credit risk.*

Data Availability: *Data are available from public sources indicated in the text.*

This paper has benefited from helpful discussions with Stephen Cooper, Dan Dhaliwal, Sergey Lobanov, Stephen Ryan, Jim Wahlen, and two anonymous reviewers. We also thank workshop participants at the Bank for International Settlements, Brigham Young University, Harvard Business School, Hebrew University, The Interdisciplinary Center at Herzliya, London Business School, Midwest Summer Research Conference at Notre Dame, University of California, Los Angeles, The University of Tennessee, and Tel Aviv University for insightful comments and suggestions.

Editor's note: This paper was accepted by Dan Dhaliwal.

Submitted August 2006
Accepted October 2007

I. INTRODUCTION

This study tests whether equity value changes associated with credit risk changes are attenuated by debt value changes associated with the credit risk changes, as Merton (1974) predicts.[1] Prior research finds that increases in credit risk are associated with decreases in equity value. However, this research does not address whether the equity value change results from a single effect or is the net of two countervailing effects—a direct effect arising from change in asset value and an indirect effect arising from change in debt value. We find that increases (decreases) in equity value are associated with decreases (increases) in debt value arising from increases (decreases) in credit risk, after controlling for the direct effect on equity value of the credit risk change. Recognition in net income of changes in values of liabilities, particularly those arising from changes in credit risk, is a controversy presently facing accounting standard setters. Thus, we provide descriptive evidence on the effects on firms' net income of recognizing presently unrecognized changes in debt value. Because fair value accounting for liabilities would apply to all firms, we conduct our tests on a broad sample of primarily solvent firms.

Credit risk changes arise from unanticipated asset value changes or asset risk changes. Merton (1974) shows that equity value changes associated with credit risk changes comprise potentially countervailing direct and indirect effects. The direct effect is the one-to-one relation between asset value and equity value that exists in the absence of debt. The indirect effect is the debt value change associated with a change in asset value or asset risk. The indirect effect associated with asset value change is the portion of the asset value change absorbed by debt holders; the indirect effect associated with asset risk change is the wealth transfer between equity holders and debt holders arising from a change in asset risk.[2] The direct and indirect effects are potentially countervailing because equity value equals asset value minus debt value. The indirect effect on equity value associated with credit risk changes is the primary focus of our study.

Understanding how credit risk changes affect the values of debt and equity is relevant to the debate about using fair value accounting for liabilities. The conceptual frameworks of the Financial Accounting Standards Board (FASB) and the International Accounting Standards Board (IASB) provide for income recognition of gains and losses arising from changes in the recognized amount of debt. Thus, if debt is recognized at fair value, then the indirect effect would result in recognizing gains (losses) associated with decreases (increases) in the fair value of debt. This is counterintuitive to some and has generated controversy relating to financial reporting for liabilities. Also, assets and liabilities are accounted for using different conventions. To the extent that recognized decreases in debt value are not offset by recognized decreases in asset value, firms with increases in credit risk could recognize net gains. Some view this as anomalous, and concern about recognizing such gains is the primary reason the European Commission endorsed International Accounting Standard (IAS) No. 39 (IASB 2003) for use by European firms only after deleting the option for firms to use fair value accounting for financial liabilities. The European Commission did not delete the corresponding option for financial assets.

---

[1] We use the term credit risk to describe factors that determine the risk premium on debt. Because we assume contractual debt cash flows do not change prior to maturity, debt value decreases result from credit risk increases. The terms credit risk, default risk, firm risk, and total risk often are used interchangeably in the studies we cite. These terms largely are consistent with our usage of the term credit risk. One exception is that default risk is sometimes more narrowly construed, i.e., although firms with no outstanding debt can have credit risk, only firms with outstanding debt have default risk (see, e.g., Dhaliwal et al. 1991 and footnote 16).

[2] Throughout, the term asset risk refers to unsystematic risk. Unanticipated changes in systematic asset risk change asset value, which is reflected in the direct effect.

To address our research question, our tests focus on the relation between annual equity returns and changes in debt value. Our proxy for change in debt value is the interaction between credit risk change and the amount of the firm's debt. Our proxy for credit risk change is change in estimated credit rating because credit ratings reflect credit rating agencies' assessments of credit risk; higher credit ratings reflect higher risk. Our proxy for the amount of debt is book value of debt, scaled by total assets. Based on Merton (1974), we predict a positive relation between equity returns and the interaction between credit risk change and the amount of debt. That is, we predict that the relation between equity returns and credit risk change is less negative for firms with more debt. Our tests include controls for credit risk change, leverage, earnings, and change in earnings. We expect equity returns are negatively related to credit risk change and positively related to earnings and changes in earnings. We do not predict the sign of the relation with leverage.

Consistent with predictions, we find that the relation between equity returns and credit risk changes is significantly less negative for firms with more debt. Because of potential nonlinearities associated with the direction of credit risk change, we also estimate the relation separately for firms with credit downgrades and upgrades. We find that downgrade firms have significantly negative equity returns. Consistent with our primary findings, we also find that equity returns are significantly less negative for firms with more debt. As predicted, we find the opposite for upgrade firms, i.e., although their equity returns are significantly positive, the returns are significantly less positive for firms with more debt. Because of potential nonlinearities in the relation associated with level of credit risk, we also permit the relation to differ depending on whether the upgrade or downgrade is within investment grade, between investment grade and non-investment grade, or within non-investment grade. Consistent with our primary findings, we find that the relation between equity returns credit risk changes is significantly attenuated for firms with more debt, except for firms downgraded within investment grade and firms upgraded to investment grade.

To investigate the robustness of our findings, we conduct several additional analyses. In particular, as an alternative proxy for debt value change we use the effect on the firm's debt value of the change in market interest rate associated with the firm's credit risk change. To calculate this proxy, we use debt maturity information disclosed in the firm's financial statement footnotes. Consistent with our primary findings, we find that the gain or loss arising from the debt value change is significantly positively associated with equity returns. Also, because credit risk changes reflect asset risk changes and asset value changes, we estimate our primary equation after replacing change in credit risk with change in equity cost of capital and change in analyst earnings forecasts. Change in equity cost of capital is a proxy for change in systematic risk; change in analyst earnings forecasts is a proxy for change in expected asset cash flows. The findings indicate that our primary results are attributable to both change in systematic risk and change in asset cash flows. In particular, we find that change in equity cost of capital and change in analyst earnings forecasts each interacted with debt attenuate the relation between returns and change in equity cost of capital and change in analyst earnings forecasts.

Establishing that changes in debt value arising from changes in credit risk are associated with changes in equity value indicates that such debt value changes are a component of firms' economic income. Because meeting the objective of financial reporting requires faithful representation of firms' liabilities and economic performance (IASB 2006), our results indicate that debt value changes are candidates for inclusion in firms' accounting income. To investigate concerns about the accounting recognition of such changes, we provide descriptive evidence on the effects on firms' reported net income of recognizing changes in debt fair values. We do this by inverting the Merton (1974) model to obtain estimates of firms' asset value and debt value. Because the estimates likely contain estimation error, this evidence should be interpreted with caution.

The evidence reveals that if all unrecognized changes in debt and asset values were recognized, then upgrade firms would report higher net income and downgrade firms would report lower net income than they currently recognize. This is consistent with firms' unrecognized asset value changes exceeding their unrecognized debt value changes. As one would expect, the evidence reveals that if only unrecognized changes in debt value were recognized, then upgrade firms would recognize lower net income and downgrade firms would recognize higher net income. However, the evidence also reveals that for downgrade firms, recognized asset write-downs are larger, on average, than unrecognized gains from decreases in debt value. This evidence mitigates the concern that debt value decreases would exceed recognized contemporaneous asset value decreases. However, this is not true for all downgrade firms, which suggests some concern about anomalous income effects is not unwarranted.

The paper proceeds as follows. Section II elaborates on the study's background, motivation, and related research. Section III describes the basis for our prediction and the research design. Section IV presents the primary findings, and Section V presents results from additional analyses. Section VI presents results relating to effects on reported net income of using fair value accounting for debt, and Section VII offers concluding remarks.

II. BACKGROUND, MOTIVATION, AND RELATED RESEARCH

Risk, Debt Values, and Equity Values

A large finance literature focuses on explaining debt values, particularly the relation between debt value and credit risk. In this literature, debt valuation models typically are based on Merton's (1974) insight that equity can be viewed as a call option on the value of underlying assets with a strike price equal to the face amount of debt (see, e.g., Duffee 1996, 1998; Duffie and Singleton 1999; Huang and Huang 2003; see Bohn 2000 for a review of this literature). The debt valuation models establish a negative relation between credit risk and debt value. This relation also applies to post-contracting changes in credit risk, because debt value changes when the market interest rate commensurate with the new level of credit risk differs from the rate determined at the inception of the debt. Strong (1990) empirically investigates debt value changes associated with credit risk changes as measured by bond rating changes for a sample of 190 firms in 1983. Strong (1990) seeks to distinguish debt value changes associated with credit risk changes from those associated with market risk changes, and finds that both explain debt value changes. However, this literature does not link debt value changes and equity value changes.

Holthausen and Leftwich (1986), Hand et al. (1992), and Goh and Ederington (1993), among others, investigate equity value changes associated with bond rating change announcements. These studies generally find that debt and equity values decrease with bond rating downgrades. Such findings indicate that downgrade announcements convey net negative information to both debt and equity markets, and are consistent with an association between credit risk increases and net negative equity value effects.[3] Consistent with these findings and with Galai and Masulis (1976) and Bowman (1979), Vassalou and Xing (2004) show that a large portion of default risk is systematic and, thus, priced in equity value. These studies generally do not find significantly positive announcement equity returns for upgrades. Prior literature explains that this likely occurs because bond rating upgrades occur with a lag and, thus, equity prices reflect the information on which the upgrade is based prior to the rating change announcement (Pinches and Singleton 1978). Regardless, these studies do not attempt to distinguish the two countervailing effects on equity value associated with credit risk changes that we document.

Other studies examine how the level of default risk affects the response of equity value to unexpected earnings. In particular, Dhaliwal and Reynolds (1994) find that earnings response coefficients (ERCs) are negatively related to the level of default risk. Dhaliwal et al. (1991; hereafter, DLF) test the empirical implications of Dhaliwal and Reynolds (1994) by using the 15-year average of leverage as a proxy for default risk. For a sample of 56 firms with high leverage and 56 firms with low leverage, DLF find low leverage firms have higher ERCs. DLF confirm this inference by finding that 48 firms with zero leverage have even higher ERCs. However, neither Dhaliwal and Reynolds (1994) nor DLF examine the effects of change in default risk or the effects on equity value of debt value changes associated with credit risk changes.

Only a few studies attempt to link debt value changes associated with credit risk changes to equity value changes. Kliger and Sarig (2000; hereafter, KS) examine changes in stock and bond prices incident to Moody's 1982 adoption of finer bond rating partitions. KS find significant decreases in debt value for firms with implied downgrades, but do not find consistently significant increases for firms with implied upgrades. The significance of equity value changes in each bond rating change group depends on the specification; KS find no positive equity returns for downgrade firms when basing expected returns on a market model. Hand et al. (1990; hereafter, HHS) investigate whether bond holders gain at the expense of stock holders when firms defease debt in substance, but not legally, for a sample of 80 defeasances by 68 firms from 1981 to 1987. For a subsample of these firms with announcement data, HHS find significantly positive bond returns and significantly negative stock returns at the defeasance announcement. However, the negative correlation between the bond returns and stock returns is weak. After also investigating motivations for the defeasances, HHS conclude that the negative stock returns are more likely attributable to information effects than to increases in debt values. These studies provide suggestive results. However, the uniqueness of the settings and small sample sizes limit generalizability of their inferences. Also, the studies do not attempt to distinguish the countervailing effect on equity value of debt value changes associated with credit risk changes.[4] Thus, the empirical validity of Merton's (1974) predictions of equity value increases associated with debt value decreases remains largely unexplored.

---

[3] Goh and Ederington (1993; hereafter, GE) find no significantly negative equity returns for bond rating downgrades associated with change in leverage. GE predict positive equity returns to these announcements because such downgrades should be associated with a wealth transfer from debt holders to equity holders. Finding no significant announcement equity returns could be because such downgrades do not reflect information that affects equity values or the return window does not span the date that the information was impounded in share prices. Our use of annual returns windows mitigates the possibility of failure to capture the effect on equity returns of credit risk changes. Also, see Ederington and Yawitz (1987) for a review of earlier studies in this literature. Relatedly, Ederington and Goh (1998) find that analysts decrease earnings forecasts following downgrades and attribute this finding to information transfer from debt rating agencies to equity analysts.

Accounting for Debt

Currently, financial accounting standards require that liabilities are initially measured at cost, which typically equals their value at the time, and subsequently measured at cost or amortized cost. Long-term debt, in particular, is initially measured at the face value of the debt adjusted for issue premium or discount, and subsequently measured at amortized cost. Interest expense is based on the rate of interest implicit in the debt issuance price. This rate reflects the debt's market rate of interest, including the effect of the firm's credit risk, when it is issued. Fair value accounting for debt would initially measure debt at the same amount. However, fair value accounting would subsequently measure debt at fair value, with changes in fair value recognized in income. Thus, the recognized amount of debt and its interest expense would reflect a current market interest rate, including the effect of the firm's current credit risk (Barth and Landsman 1995). The promised stream of cash flows associated with the debt is the same regardless of the accounting. The primary differences are in the timing of recognition of the associated cash flows, as is true of all accruals, and their characterization as principal, interest, or gains and losses.

Fair value measurement of liabilities, particularly long-term debt, is a controversy currently facing standard setters. The FASB has identified fair value as the most relevant measurement basis for financial instruments and has indicated that measuring all financial instruments at fair value is one of its long-term goals (FASB 1999). Fair value is permitted in U.S. and international standards for many financial assets, e.g., Statement of Financial Accounting Standards (SFAS) No. 133 and IAS No. 39. However, using fair value accounting for liabilities is not widespread. SFAS No. 133 and IAS No. 39 require fair value accounting for derivative liabilities, but require or permit firms to measure at amortized cost other liabilities, including long-term debt. Those who agree with the FASB's goal believe measuring liabilities at fair value is consistent with measuring assets at fair value. The FASB's goal also is consistent with income and its volatility better reflecting market and other risks when firms measure financial assets and liabilities at fair value (see, e.g., Barth et al. 1995; Hodder et al. 2006).

However, others find recognizing changes in debt value disturbing. They are particularly concerned about the financial reporting implications of recognizing changes in debt value arising from changes in the firm's own credit risk. For example, the European Central Bank called the recognition of gains associated with increases in credit risk "counterintuitive" (European Central Bank 2001), and the European Commission's endorsement of IAS No. 39 for use by European firms eliminated the fair value option for financial liabilities. The main issue of concern is that net income will not reflect changes in net asset value if debt value decreases are recognized but all concurrent asset value decreases are not. For example, if some intangible assets are not recognized, then troubled firms could report positive net income because of debt value decreases in periods in which they experience equity value decreases.[5]

---

[4] Another stream of research links debt value and equity value by simulating the potential magnitude of agency costs arising from risk-taking incentives identified in Merton (1974), e.g., Parrino and Weisbach (1999; hereafter, PW). Analytical models and simulations such as those in PW suggest agency costs can be substantial. However, PW construct simulated debt values based on the Merton (1974) model. Thus, PW cannot test the model's predictions. Empirically, Odders-White and Ready (2006; hereafter, OR) find that firms with higher credit risk have higher adverse selection equity spread components, suggesting agency costs are priced in equity value as well as in debt value. However, OR do not test the relation between debt value changes and equity value changes associated with increases in credit risk.

Lipe (2002) demonstrates how accounting ratios might convey misleading positive signals when a firm approaching bankruptcy uses fair value to measure liabilities. Lipe (2002) concludes that debt value changes attributable to credit risk changes should not be recognized. However, the misleading financial statement effects of Lipe's (2002) example primarily derive from incomplete recognition of assets and asset value changes, not from recognizing debt value changes.

Although there is a substantial literature addressing the value relevance of fair values for equity prices and returns, few studies examine fair values of liabilities in industries other than banking and insurance. Banking industry studies consistently demonstrate the value relevance of asset fair values and, to a lesser extent, deposit liabilities and long-term debt fair values. For example, Barth et al. (1996; hereafter, BBL) find that unrealized gains and losses on long-term debt estimated from disclosures in bank financial statements are significantly associated with the difference between equity market value and equity book value, although not in all model specifications and years. BBL find no significant association between changes in unrecognized gains and losses on long-term debt and changes in the difference between equity market value and equity book value. Eccher et al. (1996) and Nelson (1996) find no association between equity value and disclosed long-term debt fair value. For a sample of nonfinancial firms, Simko (1999) finds that disclosed liability fair values are associated with equity values, but not consistently across industries and years.

Barth, Landsman, and Rendleman (1998 hereafter, BLR) estimate debt values for a sample of nonfinancial firms and investigate the financial statement effects of fair value accounting for debt. BLR find that financial statement amounts based on fair value are potentially relevant to investors because financial statement amounts would be substantially different from those currently recognized. However, none of these studies investigates the effects of changes in credit risk on the values of the firm's debt and equity.

III. BASIS FOR PREDICTION AND RESEARCH DESIGN

Based on Merton (1974), equity value, E, can be expressed as:

$$E = A\mathrm{N}(d_1) - Ke^{-rT}\mathrm{N}(d_2) \tag{1}$$

where A is asset value, K is contractual debt payments, e is the exponential function, r is the risk-free rate of return, T is duration of the debt, N is the cumulative standard normal distribution, $d_1 = \frac{\ln(A/K) + (r + \sigma^2/2)^T}{\sigma\sqrt{T}}$, $d_2 = d_1 - \sigma\sqrt{T}$, and σ is standard deviation of asset returns. $N(d_1)$ reflects the probability that asset value will exceed the contractual debt payments when they become due.

---

[5] When credit risk decreases, asset value increases often are not recognized. Yet, if debt is measured at fair value, debt value increases would be recognized as losses. When credit risk increases, the opposite occurs. Recognized assets may be written down. However, not all assets are recognized and asset write-downs may not be complete or timely. Even in the absence of a change in asset value, there could be a wealth transfer from debt holders to equity holders when asset risk changes. Although such wealth transfers are consistent with Merton's (1974) theory, they are viewed by some as anomalous. This is because recognizing in net income the effects of such wealth transfers would result in recognizing a gain (loss) from decreases (increases) in debt value associated with increases (decreases) in risk. Our descriptive evidence in Section VI suggests that for most firms the "anomaly" associated with unrecognized asset value changes dominates that associated with only wealth transfers. Income effects associated with credit risk increases are more troublesome to regulators than those associated with credit risk decreases.

To isolate the indirect effect that is the focus of our study, we restate Equation (1) as follows.

$$E = A - Ke^{-rT}\{L\ (1 - N(d_1)) + N(d_2)\} \tag{2}$$

where $L = A/Ke^{-rT}$. As does Merton (1974), Equation (2) states equity value, E, as asset value, A, minus debt value, $D = Ke^{-rT}\{L\ (1 - N(d_1)) + N(d_2)\}$. Equation (2) reveals that $D = A$ when asset value is expected to be less than the present value of the contractual debt payments, and equals the present value of the contractual debt payments otherwise.

Our interest is in the relation between equity value changes and debt value changes arising from credit risk changes. Credit risk changes derive primarily from unanticipated changes in asset value or asset risk.[6] Thus, we restate Equation (2) in change form, which, holding constant the amount of debt, Ke^{-rT}, results in Equation (3) where Δ denotes change.

$$\Delta E = \Delta A - Ke^{-rT}\Delta\{L\ (1 - N(d_1)) + N(d_2)\}. \tag{3}$$

Equation (3) reveals ΔA as the direct equity value effect of an asset value change. This is the one-to-one mapping between asset value and equity value that exists in the absence of debt. Equation (3) also reveals that when the firm has debt, asset value change and asset risk change affect debt value, which results in the indirect equity value effect. This is the asset value change absorbed by debt holders and the debt value effect associated with asset risk change. Debt value changes when asset value changes even for solvent firms. Because priority of debt over equity at liquidation of the firm does not imply debt holders have a priority claim on asset value before liquidation, debt holders participate in asset value changes, even when asset value exceeds the amount of the debt.

The relation between changes in asset value and asset risk and change in debt value is complex. For asset value changes, $\partial D/\partial A > 0$. Because equity value cannot be negative, $\partial E/\partial A$ is net positive. However, Equation (3) makes clear that $\partial E/\partial A$ has two components, the positive direct effect, and the negative indirect effect stemming from $\partial D/\partial A$. For asset risk changes, $\partial D/\partial\sigma < 0$. There is no direct equity value effect of change in unsystematic risk. However, because equity value equals asset value minus debt value, asset risk increases (decreases) result in equity value increases (decreases) associated with debt value decreases (increases), even when asset value is unchanged. That is, $\partial D/\partial\sigma < 0$ means that $\partial E/\partial\sigma > 0$.

Thus, Merton (1974) predicts that equity value changes associated with credit risk changes can be characterized into potentially countervailing direct and indirect effects. Our research question leads us to focus on testing the prediction that the decrease (increase) in equity value associated with an increase (decrease) in credit risk is mitigated by debt. We test the prediction because, as with any model, the Merton (1974) model is based on assumptions that may not hold empirically. For example, the model does not consider the effects of institutional features such as market inefficiencies or debt covenants, either of which could limit equity value increases associated with debt value decreases. Also, Merton

---

[6] Change in credit risk also can arise from change in the amount of debt. Equation (4) below assumes the amount of debt does not change during the year. However, see footnote 8 for findings relating to alternative measures of debt.

(1974) shows that the indirect effect decreases as solvency increases, raising the possibility that the effect is negligible for most firms. We test the importance of the indirect effect for a broad sample of primarily solvent firms. Failure to find that the indirect effect is important would call into question the substance of the accounting controversy over recognizing the effect in net income.

Research Design: Returns and Credit Risk Changes

Our empirical specification of Equation (3) is Equation (4):

$$\begin{aligned} RET_t = {} & \beta_0 + \beta_1 \Delta CR_t + \beta_2 \Delta CR_t \times DBTA_t + \beta_3 DBTA_t + \beta_4 EPS_t + \beta_5 \Delta EPS_t \\ & + \beta_6 NEG_t + \beta_7 NEG_t \times EPS_t + \beta_8 NEG_t \times \Delta EPS_t + \varepsilon_{4t}. \end{aligned} \tag{4}$$

RET is annual size-adjusted stock return, inclusive of dividends; it corresponds to ΔE in Equation (3). t denotes year; we omit firm subscripts. For our research question, the key variable in Equation (4) is $\Delta CR \times DBTA$, where *CR* is credit risk and *DBTA* is the debt-to-assets ratio. It is our proxy for change in debt value associated with change in credit risk. Thus, estimating its coefficient permits us to test our predictions relating to the indirect equity value effect. We predict β_2 is positive.

Change in credit risk reflects change in asset value and change in asset risk. Both alter the distribution of expected debt payments, resulting in debt value changes. Thus, we use change in credit risk to capture $\Delta\{L(1 - N(d_1)) + N(d_2)\}$ in Equation (3). Our proxy for change in credit risk is ΔCR, where *CR* is a categorical variable that ranges from 1 denoting low risk to 4 denoting high risk; ΔCR is positive (negative) when credit risk increases (decreases). The Appendix explains how we follow prior research to estimate *CR* based on the relation between accounting variables and credit ratings for firms with such ratings. These variables are total assets, return on assets, ratio of debt to total assets, and indicators for whether a firm pays dividends, has subordinated debt, or has negative net income. We use credit ratings because they reflect the credit agency's assessment of credit risk, where that assessment is based on publicly available and private information (Jorion et al. 2005). Using *CR* permits us to expand our sample beyond firms with credit ratings, thereby enhancing the generalizability of our inferences. Also, as Section II notes, credit ratings, especially upgrades, are revised with a lag (e.g., Pinches and Singleton 1978), which adds noise when actual credit rating changes are used as proxies for changes in credit risk.[7]

Our proxy for the amount of debt in Equation (3), i.e., the present value of the contractual debt payment, Ke^{-rT}, is book value of debt. Book value of debt is not a perfect proxy, however, because the payments are discounted at the firm's borrowing rate when the debt was issued, rather than the current risk-free rate. *DBTA* is the end-of-year ratio of book value of debt to book value of total assets. We deflate book value of debt by total assets to control for cross-sectional size differences in Equation (4) and because doing so results in a variable that ranges from 0 to 1, which facilitates interpretation of its coefficient.[8]

---

[7] Our inferences are unaffected by the use of alternative measures of credit risk change, including change in actual credit rating. See Table 6 and footnote 21. Hillegeist et al. (2004; hereafter, HKCL) estimate probability of default using asset value and asset risk estimates obtained from inverting the Merton (1974) model. Because HKCL use stock prices as inputs, the HKCL probability of default estimates are endogenous in Equation (4). Thus, we do not use them in our tests. In Section VI, we present findings from estimating a relation analogous to Equation (4) using asset value and asset risk estimates obtained from the Merton (1974) model. Our inferences from that specification are the same as those we obtain from Equation (4).

We also include ΔCR in Equation (4). We expect ΔCR primarily to be a proxy for change in asset value, ΔA. Thus, based on Equation (3) and prior research (e.g., Holthausen and Leftwich 1986), we predict β_1 is negative. However, because ΔCR reflects asset value and asset risk changes related to debt, it likely is an incomplete and noisy proxy for asset value changes related to equity. For this reason, and to facilitate comparison with the extensive accounting literature examining the relation between returns and earnings, Equation (4) includes EPS, earnings per share before extraordinary items deflated by beginning-of-year stock price, and ΔEPS. Because EPS is earnings after interest expense, both of these variables are proxies for change in asset value related to equity. We predict β_4 and β_5 are positive. We also include NEG, $NEG \times EPS$, and $NEG \times \Delta EPS$, where NEG is an indicator variable that equals 1 if EPS is negative, and 0 otherwise, to permit the relation between returns and earnings to differ for firms with negative earnings (Hayn 1995; Barth, Beaver, and Landsman 1998).[9] We predict β_6, β_7, and β_8 are negative. We include $DBTA$ in Equation (4) because we interact it with ΔCR.[10] We do not predict the sign of its coefficient, β_3.[11]

We estimate Equation (4) pooling all firms with year and industry fixed effects, defining industries following Barth, Beaver, and Landsman (1998). To mitigate the effects of influential observations, we estimate Equation (4) using Huber-M estimation, which minimizes a less rapidly increasing function of the regression residuals than OLS.[12]

Because d_1 is a logarithmic function of asset value, A, the Merton (1974) model predicts that the sensitivity of debt value, D, to change in asset value decreases as asset value increases, i.e., as credit risk decreases. Therefore, we also estimate Equation (4) separately for firms with credit risk upgrades and downgrades. To permit further potential nonlinearities associated with level of credit risk, we partition firms based on whether the upgrade or downgrade is within investment grade, between investment grade and non-investment grade, or within non-investment grade.[13]

---

[8] None of our inferences is affected if we instead deflate book value of debt by lagged market value of equity or number of shares outstanding, or use in place of $DBTA_t$ the ratio of total liabilities to total assets, the ratio of debt and capitalized leases to total assets, $DBTA_{t-1}$, or the ratio of lagged market value of debt to lagged market value of assets that we estimate in Section VI. Our inferences also are unaffected if we include $\Delta DBTA$ in our estimating equations.

[9] Untabulated findings reveal that our inferences are unaffected if we exclude the five earnings variables. The coefficient on ΔCR is -0.24 ($t = -40.08$), the coefficient on $\Delta CR \times DBTA$ is 0.21 ($t = 10.43$), and the coefficient on $DBTA$ is -0.15 ($t = -15.41$).

[10] We use only accounting-based explanatory variables in Equation (4) to avoid endogeneity associated with changes in market values. In Section VI, as a sensitivity check, we estimate a version of Equation (4) using market value estimates. See also footnote 7.

[11] Returns may be correlated with change in debt. However, change in $DBTA$ is reflected in ΔCR through Equation (A1). Also, untabulated statistics reveal that changes in $DBTA$ are close to zero for most firms—the upper (lower) decile of $\Delta DBTA$ is 0.10 (-0.07). Inferences from an untabulated estimation of Equation (4) eliminating observations with the highest and lowest 1 percent of $\Delta DBTA$ are the same as those from the estimation we tabulate.

[12] All of our inferences are unaffected if we use OLS estimation. We also obtain similar inferences if we base our test statistics on standard errors that are clustered by firm, which controls for heteroscedasticity and intertemporal firm-specific dependence in regression residuals.

[13] This approach effectively controls for beginning-of-year credit risk, as well as clientele effects. Nonetheless, in untabulated analyses, we include CR_{t-1} as an additional explanatory variable in our estimation equations. None of our inferences is affected.

IV. DATA AND FINDINGS FOR RELATION BETWEEN CHANGES IN RISK AND EQUITY RETURNS

Data and Sample

Fair value accounting for liabilities would apply to all firms, regardless of financial condition. Thus, we construct our sample to comprise a broad cross-section of primarily solvent firms. In particular, we begin with all firms with available data on Compustat for 1986–2003.[14] We eliminate firms in the utilities, financial services, and real estate industries because their capital structures markedly differ from those of other firms. To mitigate the effects of outliers, we eliminate firms for which the absolute value of EPS_t, EPS_{t-1}, or ΔEPS_t is greater than 1.5 (Easton and Harris 1991) and firms with *RET* in the extreme top and bottom percentiles of the observations (Kothari and Zimmerman 1995; Collins et al. 1997; Fama and French 1998; Barth et al. 1999; among others). To mitigate undue effects of very small firms, we also eliminate firms with total assets or sales less than $10 million, or share price less than $1.[15] The final sample comprises 49,081 firm-year observations, of which 11,399 have credit ratings. Data limitations reduce the sample size for some analyses.

We obtain stock market data from CRSP, analyst forecast data from I/B/E/S, bond covenant data from the Fixed Income Securities Database, interest rate data from The Federal Reserve, and all other data from Compustat. The credit ratings on Compustat are Standard & Poor's Issuer Credit Rating.[16] *RET* is each firm's size-adjusted annual buy-and-hold return, computed as the firm's compounded monthly fiscal year return minus the corresponding compounded size decile return associated with the firm's market value of equity at the beginning of the year.[17]

Descriptive Statistics

Table 1 presents descriptive statistics for the variables in Equation (4). Panel A presents distributional statistics, Panel B presents correlations between the variables, and Panel C presents the industry composition of the sample. Panel A reveals that the mean of *RET* is −0.01, which is close to zero as expected. The median is somewhat more negative, −0.07, which indicates skewness in returns similar to that observed for all firms on CRSP during

---

[14] The sample period begins in 1986 because Compustat does not include credit ratings before 1985 and two years are necessary to calculate change in credit risk.

[15] Our inferences are unaffected if we include all firms in our tests.

[16] Prior to September 1, 1998, the S&P rating is the firm's senior debt rating, which is an assessment of the creditworthiness of the firm's long-term debt that is not subordinate to any other long-term debt. Typically, the rating is for the firm's most senior debt issue. If a firm does not have senior debt, it is an implied senior rating. Beginning September 1, 1998, S&P credit ratings reflect the firm's overall creditworthiness; our inferences are the same before and after September 1, 1998. S&P does not require the rated firm to have debt outstanding. Of the 11,399 observations for 1,851 firms in our sample with S&P credit ratings, 67 observations for 45 firms have no debt. Thus, we do not restrict our expanded sample to firms with debt outstanding. Of the 37,682 observations for which we estimate credit risk, 6,339 have zero debt. Our credit risk estimation procedure is applicable to firms with zero debt—four of the six variables in Equation (A1) in the Appendix do not require the firm to have debt. Also, our inferences are unaffected by (1) eliminating firms with zero debt, (2) permitting the coefficients in Equation (A1) to vary for zero debt firms, and (3) placing all zero debt firms into the lowest *DBTA* portfolio in the ranked *DBTA* specification. Our inferences also are unchanged when we eliminate observations with negative equity book value.

[17] Consistent with Fama and French (1992) and Jegadeesh (1992), our inferences are unaffected by using beta-adjusted returns. However, using beta-adjusted returns noticeably reduces our sample size. We use annual returns for our tests because there is no basis on which to identify shorter return windows for sample firms without announced credit rating changes. Even for firms with such announcements, using annual returns mitigates the potential for mis-specifying when share prices reflect the change in economic fundamentals associated with the credit rating change (Finches and Singleton 1978; Dichev and Piotroski 2001), mis-specifying market expectations at the announcement date, and confounding our inferences with information effects associated with the announcement.

TABLE 1
Descriptive Statistics
(n = 49,081)

Panel A: Distributional Statistics

Variable	Mean	Median	Std. Dev.
RET	−0.01	−0.07	0.54
ΔCR	0.01	0.00	0.44
DBTA	0.21	0.18	0.20
EPS	0.02	0.05	0.18
ΔEPS	0.01	0.01	0.19
NEG	0.26		

Panel B: Pearson (above the Diagonal) and Spearman (below the Diagonal) Correlations

	RET	*ΔCR*	*DBTA*	*EPS*	*ΔEPS*	*NEG*
RET		−0.20	−0.07	0.27	0.25	−0.25
ΔCR	−0.21		0.08	−0.21	−0.28	0.26
DBTA	−0.06	0.07		−0.13	−0.05	0.12
EPS	0.43	−0.26	−0.03		0.52	−0.65
ΔEPS	0.39	−0.37	−0.04	0.50		−0.29
NEG	−0.32	0.26	0.07	−0.76	−0.38	

Panel C: Industry Composition of Sample

Industry	SIC Codes	n	Percent
Mining and construction	1000–1999, except 1300–1399	1,266	2.58
Food	2000–2111	1,531	3.12
Textiles, printing, and publishing	2200–2799	3,871	7.89
Chemicals	2800–2824, and 2840–2899	1,618	3.30
Pharmaceuticals	2830–2836	1,600	3.26
Extractive industries	2900–2999, and 1300–1399	2,199	4.48
Durable manufacturers	3000–3999, except 3570–3579, 3670–3679	14,610	29.77
Computers	7370–7379, 3570–3579, and 3670–3679	7,039	14.34
Transportation	4000–4899	3,107	6.33
Retail	5000–5999	7,172	14.61
Services	7000–8999, except 7370–7379	5,068	10.32
		49,081	100.00

All correlations in Panel B are significantly different from 0. Sample of 7,561 Compustat firms from 1986–2003.

Variable Definitions:
RET = size-adjusted annual stock return, including dividends (from CRSP);
CR = credit risk group (1 = highest to 4 = lowest);
DBTA = ratio of debt (Compustat #9 + #44) to total assets (#6);
EPS = earnings per share before extraordinary items (#18/shares outstanding from CRSP), deflated by beginning of year stock price;
NEG = indicator for negative *EPS*; and
Δ = annual change.

the same time period. Panel A also reveals that sample firms have positive mean and median *EPS* and *ΔEPS*. Although the median credit risk change, *ΔCR*, is zero, the mean is positive, which indicates that, on average, credit risk increases. Table 1, Panel A, also reveals that mean *DBTA* is 21 percent and 26 percent of the sample firms have negative *EPS*.

Panel B of Table 1 reveals that *RET* is negatively correlated with ΔCR (Pearson correlation is −0.20 and Spearman correlation is −0.21), which is consistent with prior research and with ΔCR reflecting changes in asset value. As expected based on prior research, *RET* is positively correlated with *EPS* and ΔEPS. *RET* is negatively correlated with *DBTA*. Other correlations in Panel B also are consistent with expectations. For example, the correlations between *EPS* and ΔEPS are positive, and those between ΔCR and *EPS* and ΔEPS are negative. The negative correlation between ΔCR and *EPS* (ΔEPS) is consistent with a positive association between asset value changes and *EPS* (ΔEPS), which is consistent with earnings and credit risk change reflecting some common components of asset value change. ΔCR and *DBTA* are positively correlated, which is consistent with downgrade firms having more debt. *NEG* is negatively correlated with *RET*, *EPS*, and ΔEPS, and positively correlated with ΔCR and *DBTA*. All correlations are significantly different from zero.[18] However, we base our inferences on the Equation (4) multivariate relation.

Table 1, Panel C, reveals that the industries most highly represented in the sample are Durable Manufacturers (29.77 percent), Retail (14.61 percent), and Computers (14.34 percent). These percentages reflect the industry composition of the Compustat population. Untabulated statistics reveal that the industry composition of firms with credit ratings is similar to that in Panel C.

Primary Findings

Table 2 presents regression summary statistics from estimating Equation (4). It reveals, as predicted, that the relation between change in credit risk and equity returns is less negative for firms with more debt. In particular, the first set of columns in Panel A reveals that the coefficient on $\Delta CR \times DBTA$ is significantly positive (coef. = 0.19, t = 10.24). That is, the indirect equity value effect associated with a change in credit risk is significant. Also as predicted, the coefficient on ΔCR is significantly negative (coef. = −0.11, t = −19.56), those on *EPS* and ΔEPS are significantly positive (coefs. = 1.89 and 0.43, t = 72.11 and 26.03), and those on *NEG*, $NEG \times EPS$, and $NEG \times \Delta EPS$ are significantly negative (coefs. = −0.12, −1.80, and −0.28, t = −23.43, −57.10, and −13.19). Panel A also reveals that the coefficient on *DBTA* is significantly negative (coef. = −0.07, t = −7.82).

To compare the magnitudes of the direct and indirect effects, the second set of columns in Panel A presents summary statistics from estimating a version of Equation (4) using a ranked *DBTA* variable. The ranked *DBTA* variable is the decile rank of *DBTA*, scaled to be between 0 and 1. Specifically, we place firms into portfolios 0 to 9 with portfolio 9 comprising firms with the largest *DBTA*, and divide these portfolio ranks by 9. This permits us to interpret β_1 as the magnitude of the relation between change in credit risk and equity returns for firms with the lowest *DBTA*. The sum of β_1 and β_2 is the magnitude for firms with the highest *DBTA*.

Results of the rank regression in the second set of columns in Table 2, Panel A are consistent with predictions and the results in the first set of columns. Most importantly, the coefficient on $\Delta CR \times DBTA$ is positive, 0.12, and significantly different from zero (t = 9.75). This finding indicates that changes in debt value associated with changes in credit risk are significantly negatively associated with equity value. The sum of β_1 and β_2, −0.13

---

[18] We use the term significance to denote statistical significance at less than the 0.05 level based on a one-sided test when we have signed predictions and a two-sided test otherwise.

TABLE 2
Regression of Returns on Debt Interacted with Credit Risk Change
(n = 49,081)

$$RET_t = \beta_0 + \beta_1 \Delta CR_t + \beta_2 \Delta CR_t \times DBTA_t + \beta_3 DBTA_t + \beta_4 EPS_t + \beta_5 \Delta EPS_t + \beta_6 NEG_t + \beta_7 NEG_t \times EPS_t + \beta_8 NEG_t \times \Delta EPS_t + \varepsilon_{4t}$$

Panel A: Pooled Credit Risk Effects

				DBTA Ranks	
	Pred.	**Coef.**	**t-statistic**	**Coef.**	**t-statistic**
ΔCR	−	−0.11	−19.56	−0.13	−18.02
$\Delta CR \times DBTA$	+	0.19	10.24	0.12	9.75
$DBTA$	?	−0.07	−7.82	−0.04	−7.32
EPS	+	1.89	72.11	1.90	72.25
ΔEPS	+	0.43	26.03	0.43	26.02
NEG	−	−0.12	−23.43	−0.12	−23.59
$NEG \times EPS$	−	−1.80	−57.10	−1.80	−57.12
$NEG \times \Delta EPS$	−	−0.28	−13.19	−0.29	−13.31
Adj. R^2		0.17		0.17	

Panel B: Separate Effects for Downgrades and Upgrades

	Pred.	**Coef.**	**t-statistic**	**Coef.**	**t-statistic**
$DN \times \Delta CR$	−	−0.09	−10.90		
$DN^{INV} \times \Delta CR$	−			−0.01	−0.19
$DN^{ACR} \times \Delta CR$	−			−0.04	−3.52
$DN^{NINV} \times \Delta CR$	−			−0.13	−12.13
$DN \times \Delta CR \times DBTA$	+	0.17	7.53		
$DN^{INV} \times \Delta CR \times DBTA$	+			0.04	0.37
$DN^{ACR} \times \Delta CR \times DBTA$	+			0.15	4.88
$DN^{NINV} \times \Delta CR \times DBTA$	+			0.19	5.64
$UP \times \Delta CR$	−	−0.13	−15.42		
$UP^{INV} \times \Delta CR$	−			−0.11	−3.93
$UP^{ACR} \times \Delta CR$	−			−0.10	−7.09
$UP^{NINV} \times \Delta CR$	−			−0.16	−14.83
$UP \times \Delta CR \times DBTA$	+	0.18	4.95		
$UP^{INV} \times \Delta CR \times DBTA$	+			0.27	2.18
$UP^{ACR} \times \Delta CR \times DBTA$	+			0.09	1.39
$UP^{NINV} \times \Delta CR \times DBTA$	+			0.21	4.70
$DBTA$	?	−0.07	−7.11	−0.07	−7.17
EPS	+	1.89	72.17	1.90	72.59
ΔEPS	+	0.42	25.40	0.42	25.20
NEG	−	−0.13	−23.67	−0.12	−22.88
$NEG \times EPS$	−	−1.81	−57.36	−1.82	−57.65
$NEG \times \Delta EPS$	−	−0.27	−12.27	−0.27	−12.44
Adj. R^2		0.17		0.17	

(*continued on next page*)

TABLE 2 (continued)

Huber M-estimates are presented, with year and industry fixed effects untabulated. Sample of 7,561 Compustat firms from 1986–2003. See Table 1 for industry composition.

Variable Definitions:

RET = size-adjusted annual stock return, including dividends (from CRSP);
CR = credit risk group (1 = highest to 4 = lowest);
DBTA = ratio of debt (Compustat #9 + #44) to total assets (#6);
DBTA ranks = decile rank of *DBTA*, scaled between 0 and 1;
EPS = earnings per share before extraordinary items (#18/shares outstanding from CRSP), deflated by beginning of year stock price;
NEG = indicator for negative net income before extraordinary items;
DN (*UP*) = indicator for credit downgrade (upgrade);
DN^{INV} (UP^{INV}) = indicator for credit downgrade (upgrade) within investment grade;
DN^{NINV} (UP^{NINV}) = indicator for credit downgrade (upgrade) within non-investment grade;
DN^{ACR} (UP^{ACR}) = indicator for credit downgrade (upgrade) across grades; and
Δ = annual change.

+ 0.12, is not significantly different from zero (t = −0.70). This finding indicates that for the highest debt firms, the increase in equity value associated with a decrease in debt value, i.e., the indirect effect, essentially offsets the decrease in equity value associated with a decrease in asset value, i.e., the direct effect. Our lowest *DBTA* decile firms have zero debt. Thus, finding that β_1 is significantly negative (t = −18.02) indicates that for zero debt firms an increase in credit risk is associated with a decrease in equity value, which reflects the direct effect on equity value of change in credit risk.

Our primary findings are based on pooled estimates of Equation (4). Although Equation (4) includes year and industry fixed effects, the relations could exhibit differences across years and industries other than mean effects. However, untabulated findings reveal that this is not the case. Separate-year estimation yields positive coefficients on $\Delta CR \times DBTA$ in all 18 years, and a cross-year Z-statistic of 12.35.[19] Separate-industry estimation yields coefficients on $\Delta CR \times DBTA$ that are positive in all 11 industries; the cross-industry Z-statistic is 3.74. All other results are consistent with those in Table 2.

The first set of columns in Table 2, Panel B, relates to estimating Equation (4) permitting the coefficients on ΔCR and $\Delta CR \times DBTA$ to vary with the sign of the credit risk change. Consistent with the findings in Panel A, the Panel B findings reveal that ΔCR is significantly negatively related to *RET* for credit downgrades, $DN \times \Delta CR$, and upgrades, $UP \times \Delta CR$ (coef. = −0.09; t = −10.90 for downgrades; coef. = −0.13; t = −15.42 for upgrades). Thus, downgrade (upgrade) firms have significantly negative (positive) incremental returns. More importantly for our research question, the findings also reveal that the relation between credit downgrades (upgrades) and returns is less negative (positive) for firms with more debt. The coefficient on $\Delta CR \times DBTA$ is significantly positive for firms with credit downgrades, $DN \times \Delta CR \times DBTA$, and upgrades, $UP \times \Delta CR \times DBTA$, (coef. = 0.17; t = 7.53 for downgrades; coef. = 0.18; t = 4.95 for upgrades). Untabulated findings based on the ranked *DBTA* specification reveal the same inferences.[20]

---

[19] $Z = (\text{mean } t)/(\text{std. deviation } t/\sqrt{N-1})$, where N is the number of years (industries) and t is the t-statistic on the estimated coefficient for each year (industry) (see White 1980; Bernard 1987).

[20] That the tabulated coefficient for downgrade firms, 0.17, is smaller than that for upgrade firms, 0.18, appears inconsistent with predictions from Merton (1974). However, findings from an untabulated ranked *DBTA* specification, which permits comparison of coefficient magnitudes, reveal that the coefficient for downgrade firms, 0.11, is larger than that for upgrade firms, 0.10.

Findings from estimating Equation (4) partitioning firms based on the type of credit risk change are in the second set of columns in Table 2, Panel B. As the Appendix notes, credit rating groups 1, 2, 3, and 4 are analogous to firms with ratings of AAA to A−, BBB+ to BBB−, BB+ to BB−, and B+ to D, respectively. The first two groups comprise investment grade ratings; the second two comprise non-investment grade. In Table 2, Panel B, $DN^{INV} = 1$ for firms downgraded from the highest credit group, $CR = 1$, to the second highest, $CR = 2$, and 0 otherwise. Thus, firms with $DN^{INV} = 1$ are downgraded within investment grade. $DN^{ACR} = 1$ for firms downgraded from group 1 or 2 to group 3 or 4, and 0 otherwise. Thus, firms with $DN^{ACR} = 1$ are downgraded from investment grade to non-investment grade. $DN^{NINV} = 1$ for firms downgraded from group 3 to group 4, and 0 otherwise. Thus, firms with $DN^{NINV} = 1$ are downgraded within non-investment grade. UP^{INV}, UP^{ACR}, and UP^{NINV} are analogously defined. $UP^{INV} = 1$ for firms upgraded within investment grade, $UP^{ACR} = 1$ for firms upgraded from non-investment grade to investment grade, and $UP^{NINV} = 1$ for firms upgraded within non-investment grade.

The findings reveal that our inferences extend to almost all levels of credit risk. The coefficients on $\Delta CR \times DBTA$ interacted with DN^{ACR}, DN^{NINV}, UP^{INV}, and UP^{NINV} are significantly positive, as predicted (t-statistics range from 2.18 to 5.64). The coefficients on $DN^{INV} \times \Delta CR \times DBTA$ and $UP^{ACR} \times \Delta CR \times DBTA$ are positive, as predicted, but not significantly so (t = 0.37 and 1.39). These findings indicate that for all levels of credit risk changes, except downgrades within investment grade and upgrades to investment grade, equity value changes are significantly associated with debt value changes.[21]

V. ADDITIONAL ANALYSES

Estimates of Change in Debt Value Based on Change in Interest Rates

The findings in Table 2 reveal that the relation between returns and credit risk changes depends on the amount of debt. To provide additional evidence on whether the effect we document in Table 2 is attributable to change in debt value rather than to other effects, we calculate the gain or loss arising from the change in each firm's debt value attributable to a change in the firm's credit risk. We calculate the gain or loss using Equation (5) and then evaluate the relation between it and equity return:

$$\begin{aligned} GL\_ACC = \sum_{t=1}^{5} - DEBTt \times \left[\frac{1}{(1 + R_{end})^{t}} - \frac{1}{(1 - R_{beg})^{t}}\right] \\ - DEBT6^{+} \times \left[\frac{1}{(1 + R_{end})^{10}} - \frac{1}{(1 + R_{beg})^{10}}\right]. \end{aligned} \tag{5}$$

$DEBTt$ is debt maturing in each of the next five years and $DEBT6^{+}$ is debt maturing in six years and beyond. We multiply $DEBT6^{+}$ by factors taken to the 10th power because we assume that any debt not maturing in the next five years matures in ten years.[22] We obtain debt maturities from financial statement footnotes. The terms in brackets capture the interest rate change related to the firm's change in credit risk. To construct these terms, we use the average historical interest rate over our sample period associated with each credit risk category.[23] R_{beg} (R_{end}) is the average historical interest rate associated with the firm's credit risk, *CR*, at the beginning (end) of the year. Thus, R_{beg} differs from R_{end} only if the firm's credit risk changed during the year. Using the average rate avoids confounding temporal effects and helps insulate *GL_ACC* from effects associated with changes in market interest rates between the beginning and end of the year.

---

[21] When we limit the sample to firms without credit ratings, our inferences are unaffected. When we use actual credit ratings for firms that have them and CR for firms without them, our inferences are unaffected, except that the coefficient on $DN^{INV} \times \Delta CR \times DBTA$ is significantly positive as predicted (t = 1.79), and the coefficient on $UP^{INV} \times \Delta CR \times DBTA$ is not significantly different from zero (t = 0.78).

[22] Lack of data precludes us from making a more refined estimate. However, we alternatively assumed debt matures in six years or 20 years, with no change in our inferences. These findings suggest it is unlikely that our inferences are affected by the maturity assumption.

For comparison, we calculate the debt gain or loss implied by the parameters estimated in Equation (4). The market-implied gain or loss on debt, *GL_MKT*, is $0.19 \times \Delta CR \times DBTA$. We use 0.19 because it is the estimate of β_2 in Table 2, Panel A. *GL_ACC* and *GL_MKT* are per share, deflated by beginning of year stock price. The reduced sample size of 28,837 observations for this test reflects lack of yearly debt payment data and elimination of firms with *GL_ACC* or *GL_MKT* in the extreme top and bottom percentiles of observations, which mitigates the effect of outliers on our inferences. Table 3, Panel A, reveals that the means (standard deviations) of *GL_ACC* and *GL_MKT* are 0.00 and 0.00 (0.01 and 0.01). Panel B reveals that *GL_ACC* and *GL_MKT* are significantly positively correlated; the Pearson (Spearman) correlation is 0.81 (0.99). Correlations between *GL_ACC* and *RET* are similar to those between *GL_MKT* and *RET*.

Table 3, Panel C, presents regression summary statistics from Equation (4) with *GL_ACC* instead of $\Delta CR \times DBTA$. Inferences are the same as those we obtain from Table 2. In particular, the gain or loss on debt associated with changes in credit risk attenuates the equity gain or loss reflected in ΔCR. The coefficient on *GL_ACC* is 2.41, with a t-statistic of 7.82. All other results are similar to those in Table 2.[24]

EPS Interacted with DBTA

Equation (4) includes *EPS* and ΔEPS as proxies for change in asset value associated with change in equity value. However, change in asset value reflected in earnings could also be associated with change in debt value. Thus, we estimate Equation (6), which allows the relation between earnings and equity returns to depend on the amount of debt.

$$\begin{aligned} RET_t = {} & \beta_0 + \beta_1 \Delta CR_t + \beta_2 \Delta CR_t \times DBTA_t + \beta_{2a} EPS'_t \times DBTA_t \\ & + \beta_{2b} \Delta EPS'_t \times DBTA_t + \beta_3 DBTA_t + \beta_4 EPS'_t + \beta_5 \Delta EPS'_t \\ & + \beta_6 NEG'_t + \beta_7 NEG'_t \times EPS'_t + \beta_8 NEG'_t \times \Delta EPS'_t + \varepsilon_{6t}. \end{aligned} \tag{6}$$

EPS' is earnings per share before interest expense and extraordinary items deflated by beginning-of-year stock price, and $NEG' = 1$ if *EPS'* is negative, and 0 otherwise. Consistent with Equation (4), we predict β_{2a} and β_{2b} are negative. In this section we interpret ΔCR, EPS', and $\Delta EPS'$ as alternative proxies for change in asset value associated with change in debt value, and test for the joint explanatory power of β_2, β_{2a}, and β_{2b}.

Estimating Equation (6) allows us to relate our findings to those of Dhaliwal et al. (1991; hereafter, DLF). DLF estimate the relation between annual return and change in

---

[23] We obtain historical interest rates for the Moody's equivalent of the highest and second-highest credit risk groups from http://www.federalreserve.com. We estimate interest rates of the lowest credit risk group by adding a risk premium of 5 percent, which is the approximate spread between lower investment grade and junk bond rates during our sample period. We estimate interest rates for the remaining group by averaging those of the second-highest and lowest credit risk groups.

[24] In the spirit of Table 2, Panel B, we also estimate Equation (5) permitting different coefficients on gains and losses. Untabulated findings reveal that the coefficient on gains is significantly larger than that on losses (t = 3.36).

TABLE 3
Regression Relating Debt Gain/Loss to Returns
(n = 28,837)

Panel A: Descriptive Statistics for Regression Variables

Variable	Mean	Median	Std. Dev.
RET	−0.01	−0.07	0.53
GL_ACC	0.00	0.00	0.01
GL_MKT	0.00	0.00	0.01

Panel B: Pearson (above the Diagonal) and Spearman (below the Diagonal) Correlations

	RET	*GL_ACC*	*GL_MKT*
RET		−0.11	−0.10
GL_ACC	−0.19		0.81
GL_MKT	−0.19	0.99	

Panel C: Regression Summary Statistics from:

$$RET_t = \beta_0 + \beta_1 \Delta CR_t + \beta_2 GL\_ACC_t + \beta_3 DBTA_t + \beta_4 EPS_t + \beta_5 \Delta EPS_t + \beta_6 NEG_t + \beta_7 NEG_t \times EPS_t + \beta_8 NEG_t \times \Delta EPS_t + \varepsilon_t$$

	Pred.	Coef.	t-statistic
ΔCR	−	−0.11	−14.26
GL_ACC	+	2.41	7.82
DBTA	?	−0.09	−7.91
EPS	+	1.88	54.90
ΔEPS	+	0.50	22.89
NEG	−	−0.12	−17.26
NEG × EPS	−	−1.82	−44.22
NEG × ΔEPS	−	−0.31	−10.90
Adj. R^2		0.17	

All correlations in Panel B are significantly different from 0. Huber M-estimates are presented in Panel C, with year and industry fixed effects untabulated. Sample of 5,623 Compustat firms from 1986–2003.

$$GL\_ACC = \sum_{t=1}^{5} - DEBTt \times \left[\frac{1}{(1 + R_{end})^t} - \frac{1}{(1 + R_{beg})^t}\right] - DEBT6^{+} \times \left[\frac{1}{(1 + R_{end})^{10}} - \frac{1}{(1 + R_{beg})^{10}}\right],$$

deflated by $shares_t \times price_{t-1}$,

where:

R = interest rate associated with the firm's credit risk group, averaged over 1986 to 2003;
$DEBTt$ = debt maturing in each of the next one to five years (Compustat #44, #91–#94); and
$DEBT6^{+}$ = debt maturing in six years and beyond (#9 less the sum of #91–#94).

$$GL\_MKT = \beta \times \Delta CR \times DBTA$$

where:

β = 0.19 from Table 2, Panel A;
RET = size-adjusted annual stock return, including dividends (from CRSP);
CR = credit risk group (1 = highest to 4 = lowest);
$DBTA$ = ratio of debt (#9 + #44) to total assets (#6);
EPS = earnings per share before extraordinary items (#18/shares outstanding from CRSP), deflated by beginning of year stock price;
NEG = indicator for negative net income before extraordinary items; and
Δ = annual change.

earnings, permitting the coefficient on change in earnings, i.e., the earnings response coefficient, to vary with leverage. DLF focus on change in earnings before interest expense so that the earnings change captures asset value change related to debt and equity. DLF interpret change in earnings as unexpected earnings and leverage as a proxy for default risk. DLF predict the earnings response coefficient is smaller for high leverage firms.[25] DLF reason that if earnings provides information about the value of the firm, then the response of equity value to unexpected earnings will be affected by default risk because more of the change in asset value accrues to debt holders for firms closer to default. The first term on the right-hand side of Equation (1) captures this notion in that $\Delta E = \Delta A\ \mathrm{N}(d_1)$, where d_1 is a function of leverage.

Table 4 presents the regression summary statistics from Equation (6). It reveals that the coefficients on $\Delta CR \times DBTA$, $EPS' \times DBTA$, and $\Delta EPS' \times DBTA$ are all significantly different from zero (t = 8.07, −23.21, and 3.91), although the coefficient on $\Delta EPS' \times DBTA$

TABLE 4
Regression Allowing Earnings Response to Vary with Debt
(n = 46,025)

$$RET_t = \beta_0 + \beta_1\Delta CR_t + \beta_2\Delta CR_t \times DBTA_t + \beta_{2a}EPS'_t \times DBTA_t + \beta_{2b}\Delta EPS'_t \times DBTA_t + \beta_3 DBTA_t + \beta_4 EPS'_t + \beta_5\Delta EPS'_t + \beta_6 NEG'_t + \beta_7 NEG'_t \times EPS'_t + \beta_8 NEG'_t \times \Delta EPS'_t + \varepsilon_{6t}$$

				DBTA Ranks	
	Pred.	**Coef.**	**t-statistic**	**Coef.**	**t-statistic**
ΔCR	−	−0.12	−20.04	−0.12	−15.59
$\Delta CR \times DBTA$	+	0.16	8.07	0.07	5.40
$EPS' \times DBTA$	−	−0.80	−23.21	−0.69	−20.40
$\Delta EPS' \times DBTA$	−	0.16	3.91	0.13	4.05
$DBTA$	?	−0.22	−22.52	−0.11	−16.99
EPS'	+	0.93	41.55	1.11	35.41
$\Delta EPS'$	+	0.33	16.47	0.30	11.43
NEG'	−	−0.19	−31.12	−0.19	−31.55
$NEG' \times EPS'$	−	−0.62	−21.29	−0.62	−21.00
$NEG' \times \Delta EPS'$	−	−0.27	−12.15	−0.26	−11.54
Adj. R^2		0.15		0.14	

Huber M-estimates are presented, with year and industry fixed effects untabulated. Sample of 7,561 Compustat firms from 1986–2003. See Table 1 for industry composition.

Variable Definitions:

RET = size-adjusted annual stock return, including dividends (from CRSP);
CR = credit risk group (1 = highest to 4 = lowest);
DBTA = ratio of debt (Compustat #9 + #44) to total assets (#6);
DBTA ranks = decile rank of *DBTA*, scaled between 0 and 1;
EPS' = earnings per share before interest expense and extraordinary items [(#18 + #15)/shares outstanding from CRSP], deflated by beginning of year stock price;
NEG' = indicator for negative *EPS'*; and
Δ = annual change.

---

[25] DLF also include proxies for risk and earnings persistence. In Equation (6), ΔCR reflects risk; permitting the coefficients on EPS' and $\Delta EPS'$ to vary with NEG' allows for different persistence of positive and negative earnings.

does not have the predicted sign.[26] An untabulated F-test confirms the joint significance of the variables.[27] Untabulated separate-year and separate-industry estimations reveal coefficients on $\Delta CR \times DBTA$, $EPS' \times DBTA$, and $\Delta EPS' \times DBTA$ that are consistent with predictions in 18, 18, and 10 of 18 years, and in 11, 9, and 6 of 11 industries. The untabulated estimations also reveal Z-statistics that confirm the significance of the coefficient estimates on $\Delta CR \times DBTA$ and $EPS' \times DBTA$, and that all other findings are consistent with those in Table 4.[28]

The finding in Table 4 relating to $\Delta CR \times DBTA$ is consistent with that in Table 2, and reveals that our inferences are robust to including additional proxies for change in asset value associated with change in debt value. Although an untabulated estimation reveals that in a rank version of Equation (6) $\beta_1 + \beta_2$, which equals -0.05, is significantly negative ($t = -3.32$), this sum does not incorporate the incremental effects associated with EPS' and $\Delta EPS'$. The Table 4 findings also reveal that the effect we document is incremental to that documented in DLF. In particular, after controlling for $\Delta EPS' \times DBTA$, the coefficients on $\Delta CR \times DBTA$ and $EPS' \times DBTA$ are significantly positive and negative, respectively.

Separating Change in Risk and Change in Asset Value

Among other effects, change in credit risk reflects change in asset value. Change in asset value arises from change in expected systematic asset risk, change in expected asset cash flows, or both. In this section, we attempt to separately investigate these components of change in credit risk to determine whether our primary findings relate to both of these components or only one. In particular, we estimate Equation (7) in which we replace change in credit risk, ΔCR, with change in equity cost of capital, ΔECC, and change in analyst earnings forecasts, ΔAF:

$$\begin{aligned} RET_t = {} & \beta_0 + \beta_1\Delta ECC_t + \beta_2\Delta ECC_t \times DBTA_t + \beta_3\Delta AF_t + \beta_4\Delta AF_t \times DBTA_t \\ & + \beta_5 DBTA_t + \beta_6 EPS_t + \beta_7\Delta EPS_t + \beta_8 NEG_t + \beta_9 NEG_t \times EPS_t \\ & + \beta_{10} NEG_t \times \Delta EPS_t + \varepsilon_{7t}. \end{aligned} \tag{7}$$

We use ECC as our proxy for expected systematic asset risk because systematic equity risk is systematic asset risk after taking account of leverage. Thus, the interaction of $DBTA$ and ΔECC in Equation (7) will be significantly positive if equity returns associated with increases in expected systematic asset risk are more positive for firms with more debt. We use AF as our proxy for expected asset cash flows because asset value changes arising from changes in expected cash flows will be reflected in future earnings, which are the object of analyst forecasts.

We estimate ECC following Claus and Thomas (2001), Gebhardt et al. (2001), Gode and Mohanram (2003), and Easton (2004). Each study's estimate is based on the residual income model, after specifying a relation between equity cost of capital, equity market value, equity book value, and forecasted earnings and dividends. We use the assumptions in Dhaliwal et al. (2005). Following Dhaliwal et al. (2005) and Hail and Leuz (2006), *ECC* is the mean of these four cost of equity estimates. To mitigate the effects of error in estimating *ECC*, we eliminate observations for which $ECC < 0$ percent or $ECC > 50$ percent. *AF* is the consensus, i.e., median, one-year-ahead forecast of annual earnings per share in June of year *t*, minus the consensus two-year-ahead forecast of annual earnings per share in June of year $t-1$, scaled by stock price in June of year $t-1$. We exclude observations in the extreme top and bottom percentiles of the *AF* distribution. The reduced sample size of 22,769 observations for this analysis primarily reflects lack of analyst forecast data and our data trimming procedures.

---

[26] Untabulated findings reveal that the coefficient on $\Delta EPS' \times DBTA$ is significantly negative (-0.28, $t = -7.32$) when $EPS' \times DBTA$ is excluded from the estimating equation.

[27] Because interest expense is pretax, we implicitly assume the effective tax rate is zero. Untabulated findings reveal similar inferences when we assume either a tax rate of 35 percent or an estimate of each firm's effective tax rate, i.e., income tax expense divided by pretax income.

[28] Using EPS in Equation (6) in place of EPS' does not affect inferences relating to $\Delta CR \times DBTA$ or $EPS' \times DBTA$. However, consistent with DLF's use of $\Delta EPS'$, it affects inferences relating to $\Delta EPS' \times DBTA$. Its coefficient is significantly different from zero with an unpredicted sign (insignificantly different from zero) in the pooled (year-by-year) estimation.

Table 5 presents the findings. Panel A presents descriptive statistics and reveals that the mean and median ΔECC are -0.00. The mean (median) ΔAF is -0.02 (-0.01), which is consistent with analysts "walking down" their forecasts over time (Richardson et al. 2004). The distributional statistics for the other variables are similar to those in Table 1, Panel A. Table 5, Panel B, reveals that *RET* is negatively (positively) correlated with ΔECC (ΔAF), which is consistent with equity value decreases being correlated with systematic risk increases (expected cash flow decreases). Panel B also reveals that ΔECC (ΔAF) is significantly positively (negatively) correlated with ΔCR. Because increases in equity cost of capital and decreases in expected earnings are associated with increases in credit risk, these correlations support our use of ΔECC and ΔAF as proxies for the two components of ΔCR.

Panel C of Table 5 presents two sets of regression summary statistics from Equation (7). The first set is based on estimating Equation (7) directly. The findings from this set reveal, as expected, that ΔECC (ΔAF) is significantly negatively (positively) related to returns (t = -32.04 (10.46)). More importantly for our research question, the findings reveal that the relation is significantly less negative (positive) for firms with more debt. The coefficient on $\Delta ECC \times DBTA$ is significantly positive (t = 5.36) and that on $\Delta AF \times DBTA$ is significantly negative (t = -6.81). All other inferences are the same as those we obtain from Table 2. Untabulated summary statistics based on a *DBTA* rank regression reveal similar inferences. These findings indicate that our Table 2 findings are attributable to changes in expected systematic asset risk, as captured by ΔECC, as well as changes in expected asset cash flows, as captured by ΔAF.

The second set of findings in Panel C of Table 5 is based on estimating Equation (7) using a two-stage approach. Recall that our construction of *ECC* uses equity returns as an input. Thus, it is possible that the relation between *RET* and $\Delta ECC \times DBTA$ or ΔECC in Equation (7) stems from our construction of *ECC*, not from any economic relation. To obtain an estimate of ΔECC that does not depend on *RET*, we re-estimate Equation (7) replacing ΔECC with the predicted value from Equation (8):

$$\begin{aligned}\Delta ECC_t = {} & \beta_0 + \beta_1 \Delta CR_t + \beta_2 \Delta CR_t \times DBTA_t + \beta_3 DBTA_t + \beta_4 EPS_t \\ & + \beta_5 \Delta EPS_t + \beta_6 NEG_t + \beta_7 NEG_t \times EPS_t + \beta_8 NEG_t \times \Delta EPS_t \\ & + \beta_9 \Delta r_{f,t} + \varepsilon_{8t}. \end{aligned} \tag{8}$$

The explanatory variables in Equation (8) comprise all of those in Equation (7), which likely are related to *RET* but not to the true error in Equation (7), and the change in the risk-free interest rate, Δr_f. r_f is the annual one-year Treasury rate published on http://www.federalreserve.gov. Because the t-statistics from the second-stage estimation do not

TABLE 5
Regressions Relating Returns to Cost of Capital Changes and Analyst Earnings Forecast Revisions
(n = 22,769)

Panel A: Distributional Statistics for Regression Variables

Variable	Mean	Median	Std. Dev.
RET	−0.01	−0.05	0.49
ΔCR	0.02	0.00	0.46
ΔECC	−0.00	−0.00	0.03
ΔAF	−0.02	−0.01	0.03
DBTA	0.20	0.18	0.17
EPS	0.04	0.05	0.12
ΔEPS	0.00	0.01	0.12
NEG	0.18		

Panel B: Pearson (above the Diagonal) Spearman (below the Diagonal) Correlations

	RET	ΔCR	ΔECC	ΔAF	*DBTA*	*EPS*	ΔEPS	*NEG*
RET		−0.23	−0.31	0.16	−0.07	0.25	0.25	−0.23
ΔCR	−0.24		0.15	−0.16	0.09	−0.26	−0.34	0.32
ΔECC	−0.33	0.15		−0.07	0.04	−0.14	−0.15	0.14
ΔAF	0.22	−0.17	−0.09		−0.10	0.41	0.15	−0.45
DBTA	−0.05	0.08	0.03	−0.10		−0.07	−0.06	0.08
EPS	0.40	−0.27	−0.17	0.38	0.04		0.55	−0.64
ΔEPS	0.41	−0.41	−0.23	0.28	−0.03	0.52		−0.32
NEG	−0.28	0.32	0.12	−0.38	0.04	−0.66	−0.38	

Panel C: Regression Summary Statistics from:

$$
\begin{aligned}
RET_t = {} & \beta_0 + \beta_1 \Delta ECC_t + \beta_2 \Delta ECC_t \times DBTA_t + \beta_3 \Delta AF_t + \beta_4 \Delta AF_t \times DBTA_t \\
& + \beta_5 DBTA_t + \beta_6 EPS_t + \beta_7 \Delta EPS_t + \beta_8 NEG_t + \beta_9 NEG_t \times EPS_t \\
& + \beta_{10} NEG_t \times \Delta EPS_t + \varepsilon_{7t}
\end{aligned}
$$

				Two-Stage		
	Pred.	Coef.	t-statistic	Coef.	t-statistic	Bootstrapped t-statistic
ΔECC	−	−4.09	−32.04	−16.75	−15.27	−8.96
$\Delta ECC \times DBTA$	+	2.40	5.36	5.30	2.94	2.05
ΔAF	+	1.21	10.46	1.05	8.72	6.28
$\Delta AF \times DBTA$	−	−2.58	−6.81	−2.03	−5.13	−4.07
DBTA	?	−0.13	−8.09	−0.06	−3.59	−2.57
EPS	+	1.80	38.49	1.66	29.05	8.70
ΔEPS	+	0.80	24.37	0.40	8.10	3.40
NEG	−	−0.11	−14.28	−0.03	−2.53	−1.35
$NEG \times EPS$	−	−1.78	−29.91	−1.63	−23.85	−8.38
$NEG \times \Delta EPS$	−	−0.62	−14.15	−0.45	−9.43	−3.59
Adj. R^2		0.20		0.17		

(*continued on next page*)

TABLE 5 (continued)

Panel D: Regression Summary Statistics from:

$$\Delta ECC_t = \beta_0 + \beta_1 \Delta CR_t + \beta_2 \Delta CR_t \times DBTA_t + \beta_3 DBTA_t + \beta_4 EPS_t + \beta_5 \Delta EPS_t \\ + \beta_6 NEG_t + \beta_7 NEG_t \times EPS_t + \beta_8 NEG_t \times \Delta EPS_t + \beta_9 \Delta r_{f,t} + \varepsilon_{8t}$$

	Pred.	Coef.	t-statistic
ΔCR	+	0.01	10.56
$\Delta CR \times DBTA$	−	−0.01	−2.77
$DBTA$	?	0.01	4.06
EPS	−	−0.03	−7.93
ΔEPS	−	−0.03	−10.05
NEG	+	0.00	7.74
$NEG \times EPS$	+	0.02	5.53
$NEG \times \Delta EPS$	+	0.01	3.37
Δr_f	+	0.61	30.03
Adj. R^2		0.08	

All correlations in Panel B are significantly different from 0. Huber M-estimates are presented in Panel C, with year and industry fixed effects untabulated. In the two-stage estimation in Panel C, ΔECC is the predicted value from the Panel D regression. Bootstrapped t-statistics are based on the standard deviations of the coefficient estimates from 1,000 iterations of the two-stage analysis each based on 3,000 randomly selected sample firms, with replacement. Sample of 4,357 Compustat firms from 1986–2003.

Variable Definitions:

RET = size-adjusted annual stock return, including dividends (from CRSP);
ECC = the mean of the equity cost of capital estimates from the Gebhardt et al. (2001), Claus and Thomas (2001), Gode and Mohanram (2003), and Easton (2004) models; we calculate equity cost of capital at the end of June for each year;
CR = credit risk group (1 = highest to 4 = lowest);
AF = revision in analysts' consensus (median) earnings per share forecast, from June of prior year to June of current year (from I/B/E/S summary file), deflated by price;
$DBTA$ = ratio of debt (Compustat #9 + #44) to total assets (#6);
EPS = earnings per share before extraordinary items (#18/shares outstanding from CRSP), deflated by beginning of year stock price;
r_f = annual one-year Treasury rate; and
Δ = annual change.

reflect estimation error in the first-stage estimation, they likely are overstated. Thus, we also report bootstrapped t-statistics. Bootstrapped t-statistics are based on the standard deviations of coefficient estimates from 1,000 iterations of the two-stage analysis.[29] Inferences from the second set of findings in Panel C, which relate to the two-stage estimation, are the same as those from the first. Most importantly, the coefficient on $\Delta ECC \times DBTA$ is significantly positive (t = 2.94, bootstrapped t = 2.05) and that on $\Delta AF \times DBTA$ is significantly negative (t = −5.13, bootstrapped t = −4.07).

Panel D of Table 5 presents regression summary statistics from Equation (8), i.e., the first stage of the two-stage estimation in Panel C. The findings in Panel D are as expected. In particular, they confirm that ΔECC is significantly positively associated with ΔCR after controlling for leverage, earnings, change in earnings, and change in the risk-free interest rate (t = 10.56). The significantly negative coefficient on $\Delta CR \times DBTA$ confirms that debt attenuates this relation (t = −2.77). The significantly positive coefficient on *DBTA* indicates that leverage is positively related to equity cost of capital (t = 4.06). The remaining coefficients are consistent with expectations based on how we construct ΔECC. In particular, the significantly negative (positive) coefficients on *EPS* and ΔEPS (NEG, $NEG \times EPS$, and $\Delta NEG \times EPS$) indicate that positive earnings are associated with decreases in expected equity cost of capital (t = −7.93 and −10.05; t = 7.74, 5.53, and 3.37); and the significantly positive coefficient on Δr_f indicates that increases in expected equity cost of capital are positively associated with increases in market interest rates.

---

[29] We estimate the two-stage regression for each of 1,000 samples of 3,000 firms randomly selected with replacement from our sample firms. The bootstrapped standard error for each t-statistic is the standard deviation of each coefficient estimate across the 1,000 estimations.

Debt Covenants

Debt holders may protect the value of their debt from increases in credit risk by including covenants in debt contracts. Some covenants are designed to prevent the firm from rising above a particular level of credit risk; others are designed to minimize costs to debt holders associated with a credit risk increase. Thus, covenants can mitigate the effects of credit risk increases on debt value and, thus, on equity value (Core and Schrand 1999). To investigate this possibility, we estimate Equation (4) including ΔSP and $\Delta SP \times DBTA$ interacted with an indicator variable, *COV*, which equals 1 if more than one-half of the firm's debt issues have covenants, and 0 otherwise. The mean of *COV* for our sample is 0.53. Because *COV* is available only for firms with credit ratings, we use S&P credit ratings, *SP*, in this analysis as our proxy for credit risk. This results in a substantially smaller sample than in our primary analysis. We predict a negative coefficient on $\Delta SP \times DBTA \times COV$. We do not predict the sign of the coefficient on $\Delta SP \times COV$. Table 6 presents the findings. As predicted, $\Delta SP \times DBTA \times COV$'s coefficient is significantly negative (t = −1.93). This indicates that equity holders do not benefit as much from the presence of debt when covenants in debt contracts protect debt holders from participating in losses associated with increases in credit risk.[30]

VI. INCOME EFFECTS OF RECOGNIZING CHANGES IN DEBT VALUES

Estimating Asset and Debt Values and Asset Volatility

Merton (1974) provides a mechanism for decomposing equity value into the values of assets and liabilities. This permits us to provide descriptive evidence on the net income effects of recognizing changes in debt values. Specifically, given observed equity value and historical stock volatility, we invert the Merton (1974) model to obtain estimates of asset value and its volatility. We use these estimates to estimate debt value.

Our estimation procedure generally follows Hillegeist et al. (2004; hereafter, HKCL). Equity value is fiscal year-end share price times the number of shares outstanding. Following HKCL, we estimate equity volatility using daily stock returns and require at least 80 percent of the returns to be non-missing in the estimation period. We differ from HKCL in that we estimate the remaining term of the debt rather than assuming it equals one year. Compustat provides the amount of debt due in each of the next five years. Thus, we calculate the weighted average remaining term of the firm's debt by summing the percentage of total debt outstanding in each maturity category times the number of years in that category. Consistent with how we estimate *GL_ACC* in Section V, we assume that the

---

[30] Untabulated statistics reveal that *COV* and *DBTA* are significantly positively correlated, although the correlation is small, 0.09. To control for this, in an untabulated analysis we use *COV** in place of *COV*, where *COV** is the residual from a regression of *COV* on *DBTA*. Our inferences relating to covenants are unaffected.

TABLE 6
Regression Allowing Relation to Vary with Debt Covenants
(n = 11,399)

$$RET_t = \beta_0 + \beta_1\Delta SP_t + \beta_2\Delta SP_t \times COV_t + \beta_3\Delta SP_t \times DBTA_t + \beta_4\Delta SP_t \times DBTA_t \times COV_t + \beta_5 COV_t + \beta_6 DBTA_t + \beta_7 EPS_t + \beta_8\Delta EPS_t + \beta_9 NEG_t + \beta_{10} NEG_t \times EPS_t + \beta_{11} NEG_t \times \Delta EPS_t + \varepsilon_t$$

	Pred.	**Coef.**	**t-statistic**
ΔSP	–	−0.11	−4.99
$\Delta SP \times COV$	?	0.00	0.07
$\Delta SP \times DBTA$	+	0.13	2.88
$\Delta SP \times DBTA \times COV$	–	−0.15	−1.93
COV	?	0.00	0.73
$DBTA$	?	−0.12	−7.15
EPS	+	0.95	19.13
ΔEPS	+	0.45	14.65
NEG	–	−0.11	−11.18
$NEG \times EPS$	–	−0.80	−13.59
$NEG \times \Delta EPS$	–	−0.25	−6.08
Adj. R^2		0.16	

Huber M-estimates are presented, with year and industry fixed effects untabulated. Sample of 1,851 Compustat firms from 1986–2003.

Variable Definitions:
RET = size-adjusted annual stock return, including dividends (from CRSP);
SP = S&P credit rating (Compustat #280) (1 = highest to 4 = lowest);
DBTA = ratio of debt (#9 + #44) to total assets (#6);
EPS = earnings per share before extraordinary items (#18/shares outstanding from CRSP), deflated by beginning of year stock price;
NEG = indicator for negative net income before extraordinary items;
COV = indicator whether at least half of the outstanding debt issues have covenants (from Fixed Income Securities Database); and
Δ = annual change.

firm's remaining long-term debt is due in the 10th year. The Merton (1974) model assumes zero-coupon debt. Thus, if there are amounts due before maturity, we increase the amounts due by the amount of net interest paid in year *t*, as reported in the statement of cash flows. We also differ from HKCL in that we focus on debt rather than total liabilities and we include in assets all other liabilities. We define fair value of debt as the estimated value of the firm's assets minus the market value of equity.

For the risk-free rate, we use the annual one-year Treasury rate as in Equation (8). To mitigate effects of estimation errors, we eliminate the top and bottom percentiles of observations of each variable, except for the remaining term of the debt, the risk-free interest rate, and *DBTA*. These data requirements result in a sample of 19,118 observations. The decrease in sample size from Table 3, which also requires yearly debt maturity data, primarily results from missing net interest paid or stock return volatility data and our data trimming procedures.

Panel A of Table 7 presents descriptive statistics for inputs to and outputs from the Merton (1974) model. The mean (median) ratio of market value of equity, *MVE*, to book

TABLE 7
Descriptive Statistics using Merton Model Estimates
(n = 19,118)

Panel A: Distributional Statistics for Merton Model Estimation Inputs and Outputs

Variable	Mean	Median	Std. Dev.
MVE/BVE	2.09	1.65	1.67
MVA/BVA	1.74	1.43	1.11
MVD/BVD	1.20	1.11	0.56
Term remaining on debt	4.67	5.00	1.96
Risk-free interest rate	0.05	0.05	0.02
σ_E	0.49	0.44	0.23
σ_V	0.38	0.33	0.20
ΔMVA	0.00	0.01	0.96
DBTA	0.24	0.22	0.17

Panel B: Regression Summary Statistics from:

$$RET_t = \beta_0 + \beta_1 \Delta MVA_t + \beta_2 MVA_t \times DBTA_t + \beta_3 \Delta\sigma_{V_t} + \beta_4 \Delta\sigma_{V_t} \times DBTA_t + \beta_5 DBTA_t + \varepsilon_t$$

	Pred.	Coef.	t-statistic
ΔMVA	+	0.34	71.10
$\Delta MVA \times DBTA$	−	−0.37	−27.67
$\Delta\sigma_V$	0/−	0.03	1.01
$\Delta\sigma_V \times DBTA$	+	0.37	3.56
DBTA	?	−0.34	−22.12
Adj. R^2		0.22	

Huber M-estimates are presented, with year and industry fixed effects untabulated. Sample of 3,994 Compustat firms from 1986–2003.

Variable Definitions:

RET = size-adjusted annual stock return, including dividends (from CRSP);
MVE = market value of equity (from CRSP);
BVE = book value of equity (Compustat #60);
MVA = market value of assets estimated using the Merton (1974) model; *MVD* = *MVA* − *MVE*, *BVA* = book value of assets (#6) − book value of liabilities other than debt (#181 − #9 − #44);
BVD = book value of debt (#9 + #44);
σ_E = volatility of equity values estimated using monthly stock returns (from CRSP) over a period equal to the term remaining on debt;
σ_V = volatility of asset values estimated using the Merton (1974) model;
DBTA = ratio of debt (#9 + #44) to total assets (#6); and
Δ = annual change.

ΔMVA is deflated by MVE_{t-1}.

value of equity, *BVE*, is 2.09 (1.65), which is consistent with sample firms having unrecognized net assets. The average remaining maturity of debt is 4.67 years, the average risk-free interest rate over the sample period is 5 percent, and the average volatility of equity returns is 49 percent. Consistent with leverage increasing equity volatility, our calculated asset volatility, σ_V, is 38 percent. The mean (median) ratio of market value to book value of assets, *MVA/BVA*, is 1.74 (1.43). The mean (median) ratio of market value to book value of debt, *MVD/BVD*, is 1.20 (1.11). Untabulated statistics confirm that our sample comprises primarily solvent firms. In particular, asset value is less than the book value of debt for only 1.4 percent of the sample firms.

Internal Validity Check of Model Estimates

Before turning to descriptive evidence on the income effect of recognizing changes in debt values, we investigate the internal validity of our model estimates. We do this by estimating a version of Equation (4) in which we include change in asset value and change in asset volatility, both estimated using the Merton model, in lieu of credit risk change and earnings. Recall that Equation (4) includes credit risk change as a proxy for changes in asset value and asset risk. Thus, in this specification, we interact both with *DBTA*. If the model is validly implemented, the presence of debt should attenuate the equity value effects of changes in asset value and asset risk.[31]

Table 7, Panel B, presents the findings, which are consistent with our primary results. It reveals that equity returns are significantly positively associated with change in asset value, ΔMVA, (t = 71.10). Consistent with ΔMVA reflecting change in systematic risk, the coefficient on $\Delta\sigma_V$ is not significantly different from zero after controlling for ΔMVA (t = 1.01). More importantly for our research question, consistent with Merton (1974) and the Table 2 findings, the association between equity returns and increases in asset value (asset risk) is less (more) positive when the firm has more debt. In particular, the coefficient (t-statistic) on $\Delta MVA \times DBTA$ is −0.37 (−27.67) and that on $\Delta\sigma_V \times DBTA$ is 0.37 (3.56).

Descriptive Evidence on Income Effects

Table 8 presents descriptive statistics relating to the effect on net income of recognizing presently unrecognized change in debt value, ΔUD. Because of estimation error likely in ΔUD and ΔUA, change in unrecognized assets, these statistics should be interpreted with caution. Nonetheless, they indicate the potential effect on net income of recognizing changes in debt value. Table 8 presents statistics separately for firms with credit downgrades, upgrades, and no change in credit standing based on ΔCR. *NI* is net income as reported by the firms. All variables are deflated by beginning of year *MVE*.

Turning first to downgrade firms, Table 8, Panel A, reveals that, on average, these firms have negative *NI* in year *t*. They have positive *NI* in year *t*−1 and, thus, a negative change in *NI* from year *t*−1 to year *t*. The means (medians) of NI_t, NI_{t-1}, and ΔNI are −0.05, 0.06, and −0.12 (−0.02, 0.06, and −0.08), respectively. Table 8 also reveals that downgrade firms have negative mean (median) changes in unrecognized debt value, ΔUD, −0.01 (−0.00) and unrecognized asset value, ΔUA, −0.13 (−0.14). The net change in unrecognized asset and liability values, $\Delta UA - \Delta UD$, also is negative, with a mean (median) of −0.12 (−0.13). These statistics all are consistent with downgrade firms experiencing a decline in economic fundamentals between year *t*−1 and year *t*.

Negative ΔUD statistics for downgrade firms indicate that net income would have been higher had they recognized ΔUD. Untabulated statistics reveal both the mean and median ΔUD are significantly different from zero. $NI - \Delta UD$ is what net income would have been had unrecognized change in debt value been recognized. For downgrade firms, Table 8, Panel A, reveals that the mean (median) $NI - \Delta UD$ is −0.04 (−0.01). Thus, on average, downgrade firms would have negative net income even if unrecognized debt value were recognized.

---

[31] We view this as an internal validity check because stock price and its volatility are model inputs. Thus, the dependent variable is used indirectly to estimate the explanatory variables.

TABLE 8
Descriptive Statistics using Merton Model Estimates, Separately for Upgrades and Downgrades
(n = 19,118)

Panel A: Distributional Statistics by Change in Credit Risk

	Downgrades (n = 1,719)			Upgrades (n = 1,626)			No Change (n = 15,773)		
	Mean	Median	Std. Dev.	Mean	Median	Std. Dev.	Mean	Median	Std. Dev.
NI_t	−0.05	−0.02	0.16	0.10	0.08	0.08	0.04	0.06	0.11
NI_{t-1}	0.06	0.06	0.08	0.00	0.04	0.13	0.04	0.06	0.10
ΔNI_t	−0.12	−0.08	0.15	0.10	0.05	0.14	0.00	0.01	0.10
ΔUD_t	−0.01	−0.00	0.20	0.02	0.00	0.19	0.01	0.00	0.21
ΔUA_t	−0.13	−0.14	0.56	0.14	0.10	0.77	−0.02	−0.02	0.60
$\Delta UA_t - \Delta UD_t$	−0.12	−0.13	0.50	0.12	0.08	0.74	−0.02	−0.03	0.55
$NI_t - \Delta UD_t$	−0.04	−0.01	0.25	0.08	0.07	0.20	0.04	0.05	0.23
$\Delta NI_t - \Delta UD_t$	−0.10	−0.07	0.25	0.08	0.05	0.23	−0.00	0.00	0.23

Panel B: Comparison of Income and Pro Forma Income

		NI		*NI* − Δ*UD*	
NI Range	n	Mean	Median	Mean	Median
Downgrades (n = 1,655 of 1,719)					
−0.45 to −0.40	16	−0.43	−0.42	−0.38	−0.38
−0.40 to −0.35	20	−0.37	−0.37	−0.30	−0.36
−0.35 to −0.30	26	−0.32	−0.32	−0.24	−0.33
−0.30 to −0.25	32	−0.28	−0.28	−0.31	−0.25
−0.25 to −0.20	57	−0.22	−0.21	−0.16	−0.21
−0.20 to −0.15	80	−0.17	−0.17	−0.18	−0.17
−0.15 to −0.10	127	−0.12	−0.12	−0.12	−0.14
−0.10 to −0.05	212	−0.07	−0.07	−0.09	−0.07
−0.05 to 0.00	365	−0.02	−0.02	−0.01	−0.03
0.00 to 0.05	360	0.03	0.03	0.05	0.03
0.05 to 0.10	270	0.07	0.07	0.10	0.08
0.10 to 0.15	70	0.12	0.12	0.12	0.16
0.15 to 0.20	20	0.17	0.17	0.23	0.21
Upgrades (n = 1,603 of 1,626)					
−0.10 to −0.05	11	−0.06	−0.06	0.02	−0.07
−0.05 to 0.00	27	−0.02	−0.02	−0.05	−0.03
0.00 to 0.05	343	0.03	0.03	0.00	0.02
0.05 to 0.10	614	0.07	0.07	0.07	0.07
0.10 to 0.15	327	0.12	0.12	0.12	0.12
0.15 to 0.20	153	0.17	0.17	0.14	0.16
0.20 to 0.25	59	0.22	0.22	0.14	0.20
0.25 to 0.30	37	0.27	0.26	0.21	0.25
0.30 to 0.35	19	0.32	0.32	0.29	0.28

(continued on next page)

TABLE 8 (continued)

All Panel A means and medians are significantly different from 0, using a t-test for means or signed rank test for medians, except the median $\Delta NI_t - \Delta UD_t$ for firms without risk changes, and the median $NI_t - \Delta UD_t$ for downgrade firms.

Panel B presents statistics for income (*NI*) and pro forma income ($NI - \Delta UD$), after sorting observations into 5 percent intervals of the *NI* distribution. Intervals with less than ten observations are excluded. Sample of 3,994 Compustat firms from 1986–2003.

The market values of net assets and debt are estimated using the Merton (1974) model. Each is deflated by beginning of period market value of equity.

Variable Definitions:

NI = income before extraordinary items (Compustat #18);
UA (unrecognized assets) = market value of net assets − book value of net assets; and
UD (unrecognized debt) = market value of debt − book value of debt.

The fact that not all concurrent asset value changes are recognized contributes to the concern about potentially anomalous income effects arising from recognizing change in debt value. This should be less of a problem for downgrade firms because required asset write-downs reduce *NI*, at least for recognized assets. When assessing the net income effect of recognizing ΔUD, one would like to compare ΔUD to the amount of recognized asset write-downs, which is not available to us. However, if NI_{t-1} is a proxy for income before the decline in the firm's economic fundamentals, then ΔNI is a proxy for asset write-downs recognized in year *t*.

Table 8, Panel A, reveals that for downgrade firms, mean (median) $\Delta NI - \Delta UD$ is −0.10 (−0.07). These statistics indicate that for downgrade firms, on average, recognized asset write-downs are larger than unrecognized decreases in debt value. Thus, on average, the net effect of recognized decreases in asset value and increases in debt value is negative for these firms, which is consistent with them experiencing a decline in economic fundamentals. Untabulated statistics reveal that for approximately 73 percent (27 percent) of downgrade firms, recognized asset write-downs are larger (smaller) than unrecognized gains from decreases in debt value. This suggests the concern that debt value decreases would exceed recognized contemporaneous asset value decreases is unwarranted (warranted) for a large majority (substantial minority) of downgrade firms. Also, the negative mean and median $\Delta UA - \Delta UD$ reveals that, on average, downgrade firms have unrecognized decreases in equity value. This indicates that if ΔUD were recognized and these unrecognized decreases in asset value were not, net income would be higher than justified by the net change in value of the firms' assets and liabilities.

Relating to upgrade firms, Table 8, Panel A, reveals, on average, these firms have positive *NI* in year *t* and year $t-1$, and positive change in *NI* from year $t-1$ to year *t*. The means (medians) for NI_t, NI_{t-1}, and ΔNI are 0.10, 0.00, and 0.10 (0.08, 0.04, and 0.05), respectively. It also reveals that the mean (median) ΔUD is positive, 0.02 (0.00). As a result, had these firms recognized ΔUD, their net income would have been lower. Mean (median) ΔUA also is positive for these firms, 0.14 (0.10), as is mean (median) $\Delta UA - \Delta UD$, 0.12 (0.08). Overall, these statistics reveal a picture opposite to that of downgrade firms. That is, the statistics are consistent with upgrade firms experiencing an improvement in economic fundamentals between year $t-1$ and year *t*. Table 8, Panel A, reveals that the mean (median) $NI - \Delta UD$ is positive for upgrade firms, 0.08 (0.07). Thus, on average, upgrade firms would have positive net income even if unrecognized change in debt value were recognized. Mean (median) $\Delta NI - \Delta UD$ also is positive, 0.08 (0.05), which indicates that recognized increases in asset values exceed unrecognized increases in debt values.

Relating to firms with no change in credit risk, the statistics in Table 8 are as expected. For example, mean (median) ΔNI, ΔUD, and $\Delta NI - \Delta UD$ are small, 0.00 (0.01), 0.01 (0.00), and −0.00 (0.00), respectively. However, all variables have noticeable standard deviations, which is consistent with some of these firms experiencing larger effects or with random estimation error.

Table 8, Panel B, tabulates separately for downgrade and upgrade firms the mean and median NI and $NI - \Delta UD$ for each 5 percent of NI/MVE band for which there are at least ten observations. For downgrade (upgrade) firms NI ranges from −45 percent to 20 percent (−10 percent to 35 percent) of MVE. For downgrade firms, for most NI bands the means and medians of $NI - \Delta UD$ are somewhat less negative or more positive than NI. This is consistent with the overall statistics in Panel A that indicate these firms would report higher net income if change in debt value were recognized. However, the differences are small. The signs of mean and median NI do not differ from the signs of mean and median $NI - \Delta UD$ in any NI band, which indicates that the effect on net income of recognizing changes in debt value is not large enough to change the sign of net income for most downgrade firms. It also indicates that the differences between NI and $NI - \Delta UD$ are fairly uniform across the distribution of NI.

For upgrade firms, Panel B reveals that for most NI bands the means and medians of $NI - \Delta UD$ are somewhat more negative or less positive than NI. As with downgrade firms, this is consistent with the overall statistics in Panel A that indicate these firms would report lower net income if changes in debt value were recognized. Also as with downgrade firms, the signs of mean and median NI do not differ from the signs of mean and median $NI - \Delta UD$ in any NI band. This finding indicates that the effect on net income of recognizing changes in debt value is not large enough to change the sign of net income for most upgrade firms. It also indicates that the differences between NI and $NI - \Delta UD$ are fairly uniform across the distribution of NI.[32]

VII. CONCLUSION

This study tests whether equity value reflects gains and losses associated with changes in the value of debt, consistent with predictions of Merton (1974). It contributes not only to the extant debt and equity valuation literature, but also to the debate about using fair value accounting for liabilities. If fair values were recognized, then firms experiencing increases in credit risk would recognize gains because increases in credit risk result in decreases in debt value; the opposite would be the case for firms experiencing decreases in credit risk. These outcomes are counterintuitive to some—they contradict the views that debt holders of solvent firms are insulated from declines in the firms' economic fundamentals because debt has priority over equity, and that equity holders are the sole beneficiaries of firms' upside potential.

Consistent with prior research, we find that equity returns are significantly negatively related to changes in credit risk. More importantly for our research question, we find that the relation between credit risk change and equity returns is significantly less negative when the firm has more debt. This result is consistent with debt holders sharing in wealth increases and subsidizing wealth decreases. When we consider separately upgrade and downgrade firms, we find that equity returns for downgrade firms are significantly less negative when the firm has more debt, and we find the opposite for upgrade firms. Our findings hold for all credit risk groups, except for firms downgraded within investment grade and upgraded to investment grade. Thus, equity increases associated with increases in credit risk are evident for a broad cross-section of firms, including quite solvent firms.

---

[32] For firms with no change in credit risk, untabulated statistics reveal that NI and $NI - \Delta UD$ differ little at the mean and median for all NI bands.

As an alternative way to link equity value changes and debt value changes associated with credit risk changes, we calculate the gain or loss arising from change in debt value associated with a firm's change in credit risk and use it in our estimating equation in lieu of the credit risk change and debt interaction variable. Consistent with our primary findings, we find that the gain or loss is significantly positively associated with equity returns. We also find that the effect we document is associated with changes in systematic risk, as reflected in changes in equity cost of capital, and changes in expected cash flows, as reflected in analyst earnings forecast revisions.

Our findings link and empirically document the existence of two countervailing equity value effects associated with increases in credit risk: (1) decreases in equity value, presumably arising from decreases in asset value, and (2) increases in equity value associated with decreases in debt value, presumably arising from decreases in asset value or increases in asset risk. These findings indicate that changes in debt value are associated with predictable and measurable effects on changes in equity value.

Establishing that changes in debt value arising from changes in credit risk are associated with changes in equity value for a broad sample of primarily solvent firms indicates that such debt value changes are component of firms' economic income. Because faithful representation of firms' liabilities and income is consistent with the conceptual framework underlying financial reporting, our results indicate that debt value changes are candidates for inclusion in firms' accounting income. Thus, we provide evidence on what firms' reported net income would be if changes in debt value were recognized in order to inform the accounting debate about recognizing in net income such changes. We do this by inverting the Merton (1974) model to obtain an estimate of each firm's asset and debt value and asset volatility.

We find that upgrade firms would recognize higher net income than they do under current accounting standards if all changes in debt and asset values were recognized, and downgrade firms would recognize lower net income. This is consistent with firms' unrecognized asset value changes exceeding their unrecognized debt value changes. As one would expect, we also find that if only unrecognized changes in debt value were recognized, on average, upgrade firms would recognize lower net income and downgrade firms would recognize higher net income. However, we find that for downgrade firms recognized asset write-downs are larger, on average, than unrecognized gains from decreases in debt value, which mitigates the concern that debt value decreases would exceed recognized contemporaneous asset value decreases. Because this does not hold for all downgrade firms, the concern is not unwarranted for some firms. Our results suggest that anomalous effects on net income more likely arise from the failure to recognize all changes in asset values, than from the recognition of changes in debt values.

APPENDIX
CREDIT RISK ESTIMATION

Estimation Equation

We estimate the relation between credit rating and financial statement variables using the subsample of firms with credit ratings (Barth, Beaver, and Landsman 1998; Ashbaugh et al. 2006). We set *CR*, our proxy for credit risk, equal to the predicted value from Equation (A1) for firms with and without credit ratings. ΔCR in Equation (4) is the annual change in *CR*.

$$SP_t = a_0 + a_1 TA_t + a_2 ROA_t + a_3 DBTA_t + a_4 DIV_t + a_5 SUBDBT_t + a_6 NEG_t + \upsilon_t. \tag{A1}$$

SP_t is the firm's S&P credit rating at the end of year t; *TA* is the natural logarithm of end-of-year total assets; *ROA* is income before extraordinary items divided by total assets; and *DIV*, *SUBDBT*, and *NEG* are indicator variables that equal 1 if in year t the firm pays a cash dividend, has subordinated debt, or has negative *ROA*.[33] Estimating Equation (A1) using annual data to calculate *CR* and then calculating annual changes in *CR* for use in our tests mitigates the effects of credit ratings being revised with a lag (Pinches and Singleton 1978).[34]

We estimate Equation (A1) with year and industry fixed effects. *SP* ranges from 1 to 4, where larger *SP* corresponds to higher risk; groups 1, 2, 3, and 4 include firms with ratings of AAA to A−, BBB+ to BBB−, BB+ to BB−, and B+ to D, respectively.[35] Because *SP* has integer values, we use maximum likelihood estimation and an ordered probit model. We predict a_1, a_2, and a_4 are negative, and a_3, a_5, and a_6 are positive. We have no prediction for a_0.

Empirical Estimates

Table A1, Panel A, presents regression summary statistics from Equation (A1) for the 11,399 observations for firms with credit ratings. Consistent with prior research, S&P credit ratings, *SP*, are significantly negatively related to *TA*, *ROA*, and *DIV*, and significantly positively related to *DBTA*, *SUBDBT*, and *NEG*. The pseudo R^2 from the estimation is 0.66, indicating that these variables explain a substantial portion of the variation in credit ratings.[36]

Table A1, Panel B, presents the distributions of actual credit rating levels and changes and the distributions of estimated credit risk levels and changes. It reveals the distributions are similar. However, in group 1 there are fewer firms with estimated credit risk (10.11 percent) than with actual ratings (29.25 percent). The opposite is true for group 4, which comprises 40.24 percent of firms with estimated credit risk, but only 19.55 percent of firms with actual credit ratings. Panel B also reveals that changes in actual ratings are concentrated in the 0, 1, and −1 change groups, whereas changes in estimated credit risk are more

---

[33] Ashbaugh et al. (2006) also includes in Equation (A1) interest coverage and capital intensity. We do not include these variables because doing so noticeably reduces our sample size. However, our inferences are unchanged if we include these variables and conduct our tests on the reduced sample. Also, Equation (A1) does not include variables related to debt covenants. Thus, our estimated credit risk might not capture all aspects of debt relevant to its value. Implicitly, our design assumes debt for unrated firms has covenants similar to those of debt for rated firms. To the extent this assumption is not valid, our tests could be biased. The direction of the bias is not obvious. However, Table 6 reports results when we control for the existence of covenants our inferences are unchanged.

[34] Our inferences are unchanged if we estimate Equation (4) using two-year returns for our primary results and those based on only firms with credit ratings (see Table 6).

[35] Prior studies partition group 4 into two groups—one for credit ratings of B+ to B− and one for ratings CCC+ to D. We combine these two groups because the CCC+ to D group has few observations; these two groups combined have fewer observations than do the other three credit rating groups.

[36] Because *CR* is the estimated, rather than actual, credit rating, the standard errors from Equation (4) are biased downward. To correct for the additional variance in *CR*, as a robustness check, we add a component to the estimated variance of the parameters estimated in Equation (4). We obtain the added component from bootstrapping Equation (A1). Specifically, following Petrin and Train (2002, footnote 11), we repeatedly estimate Equations (A1) and (4) using bootstrapped samples. The added component is the variance in the Equation (4) parameter estimates obtained over the bootstrapped samples. Our inferences are unaffected by using this procedure.

TABLE A1
Credit Risk Estimation

Panel A: Regression Summary Statistics from

$$SP_t = a_0 + a_1TA_t + a_2ROA_t + a_3DBTA_t + a_4DIV_t + a_5SUBDBT_t + a_6NEG_t + v_t$$

	Pred.	Coef.	t-statistic
TA	−	−0.57	−57.99
ROA	−	−4.28	−19.10
DBTA	+	2.19	31.80
DIV	−	−1.03	−40.05
SUBDBT	+	0.35	13.96
NEG	+	0.31	8.59
Pseudo R^2		0.66	

Panel B: Distributions of Actual and Estimated Credit Rating Groups

		Actual		Estimated	
		n	Percent	n	Percent
Credit Rating Group					
AAA to A−	1	3,685	29.25	4,962	10.11
BBB+ to BBB−	2	3,298	26.18	8,793	17.92
BB+ to BB−	3	3,151	25.01	15,576	31.74
B+ to D	4	2,463	19.55	19,750	40.24
Change in Credit Rating Group					
Upgrades	−3			4	0.01
	−2	6	0.05	72	0.15
	−1	381	3.34	4,143	8.44
No change	0	10,394	91.18	40,081	81.66
Downgrades	1	585	5.13	4,673	9.52
	2	28	0.25	99	0.20
	3	5	0.04	9	0.02

Panel A is based on an ordered probit estimation using the 11,399 observations for firms with credit ratings. The model is estimated with year and industry fixed effects (untabulated). Estimated credit rating groups in Panel B are predicted values from the Panel A regression. Sample of Compustat firms from 1986–2003.

Variable Definitions:

SP = S&P credit rating (Compustat #280) (1 = highest to 4 = lowest);
TA = natural log of total assets (#6), in $ millions;
ROA = return on assets; net income before extraordinary items (#18) divided by total assets;
DBTA = ratio of debt (#9 + #44) to total assets;
DIV = 1 if the firm paid a cash dividend (#21) in year *t*, and 0 otherwise;
SUBDBT = 1 if the firm has subordinated debt (#80), and 0 otherwise; and
NEG = 1 if *ROA* is negative, and 0 otherwise.

widely distributed. These distributional differences are not unexpected because the explanatory variables reflect systematic differences between firms with and without credit ratings. For example, firms with credit ratings tend to have larger total assets.[37]

---

[37] The validity of *CR* as a proxy for credit risk does not depend on consistent levels of the prediction variables between firms with and without credit ratings. Rather, its validity depends on consistency of the parameters associated with the explanatory variables between the two groups of firms. Unfortunately, we are unable to determine this because credit ratings are not observable for firms without them. However, as Table 6 reports, our inferences are unaffected by using only firms with credit ratings.

As an internal validity check, we compare actual and estimated credit ratings for firms with actual ratings. Because for these firms *CR* is the predicted value for an observation used to estimate Equation (A1), the comparisons should be interpreted cautiously. However, untabulated statistics reveal similar distributions for actual and estimated credit ratings. For actual ratings, groups 1 through 4 are 31 percent, 27 percent, 24 percent, and 18 percent of the observations; for estimated ratings they are 33 percent, 27 percent, 23 percent, and 17 percent. The statistics also reveal that Equation (A1) correctly predicts 62 percent of actual ratings; it has prediction errors of more than one credit rating group only 2 percent of the time.

REFERENCES

Ashbaugh, H., D. Collins, and R. LaFond. 2006. The effects of corporate governance on firms' credit ratings. *Journal of Accounting and Economics* (42): 203–243.

Barth, M. E., W. R. Landsman, and J. M. Wahlen. 1995. Fair value accounting: Effects on banks' earnings volatility, regulatory capital, and value of contractual cash flows. *Journal of Banking & Finance* (19): 577–650.

———, and ———. 1995. Fundamental issues related to using fair value accounting for financial reporting. *Accounting Horizons* (9): 97–108.

———, W. H. Beaver, and W. R. Landsman. 1996. Value relevance of banks' fair value disclosures under SFAS 107. *The Accounting Review* (71): 513–538.

———, ———, and ———. 1998. Relative valuation roles of equity book value and net income as a function of financial health. *Journal of Accounting and Economics* (25): 1–34.

———, ———, and R. Rendleman. 1998. Option pricing-based value estimates and a fundamental components approach to account for corporate debt. *The Accounting Review* (73): 73–103.

———, ———, J. R. M. Hand, and W. R. Landsman. 1999. Accruals, cash flows, and equity values. *Review of Accounting Studies* (3): 205–229.

Bernard, V. L. 1987. Cross-sectional dependence and problems in inference in market-based accounting research. *Journal of Accounting Research* (25): 1–48.

Bohn, J. 2000. A survey of contingent-claims approaches to risky debt valuation. *Journal of Risk Finance* (Spring): 1–18.

Bowman, R. G. 1979. The theoretical relationship between systematic risk and financial (accounting) variables. *The Journal of Finance* (34): 617–630.

Claus, J., and J. Thomas. 2001. Equity premia as low as three percent? Evidence from analysts' earnings forecasts for domestic and international stock markets. *The Journal of Finance* (56): 1629–1666.

Collins, D. W., E. L. Maydew, and I. S. Weiss. 1997. Changes in the value-relevance of earnings and equity book values over the past forty years. *Journal of Accounting and Economics* (24): 39–67.

Core, J. E., and C. M. Schrand. 1999. The effect of accounting-based debt covenants on equity valuation. *Journal of Accounting and Economics* (27): 1–35.

Dhaliwal, D. S., K. J. Lee, and N. L. Fargher. 1991. The association between unexpected earnings and abnormal security returns in the presence of financial leverage. *Contemporary Accounting Research* (8): 20–42.

———, and S. S. Reynolds. 1994. The effect of the default risk of debt on the earnings response coefficient. *The Accounting Review* (69): 412–420.

———, S. Heitzman, and O. Z. Li. 2005. Taxes, leverage, and the cost of equity capital. *Journal of Accounting Research* (44): 691–723.

Dichev, I., and J. Piotroski. 2001. Long-run stock returns following bond rating changes. *The Journal of Finance* (56): 173–203.

Duffee, G. R. 1996. Estimating the price of default risk. Finance and Economic Discussion Series, Federal Reserve Board, 96–129. Washington, D.C.: Government Printing Office.

———. 1998. The relation between treasury yields and corporate bond yield spreads. *The Journal of Finance* (53): 2225–2241.

Duffie, D., and K. Singleton. 1999. Modeling term structures of defaultable bonds. *Review of Financial Studies* (12-4): 687–720.

Easton, P., and T. Harris. 1991. Earnings as an explanatory variable for return. *Journal of Accounting Research* (29): 19–36.

———. 2004. PE ratios, PEG ratios, and estimating the implied expected rate of return on equity capital. *The Accounting Review* (79): 73–95.

Eccher, E., K. Ramesh, and R. Thiagarajan. 1996. Fair value disclosures by bank holding companies. *Journal of Accounting and Economics* (22): 79–117.

Ederington, L. H., and J. Yawitz. 1987. The bond rating process. In *Handbook of Financial Markets and Institutions*, 6th edition, edited by E. Altman. New York, NY: John Wiley & Sons.

———, and J. C. Goh. 1998. Bond rating agencies and stock analysts: Who knows what when? *Journal of Financial and Quantitative Analysis* (33): 569–585.

European Central Bank. 2001. *Fair Value Accounting in the Banking Sector.* Frankfurt, Germany: ECB.

Fama, E. F., and K. R. French. 1992. The cross-section of expected stock returns. *The Journal of Finance* (47): 427–465.

———, and ———. 1998. Taxes, financing decisions, and firm value. *The Journal of Finance* (53): 819–843.

Financial Accounting Standards Board (FASB). 1998. *Accounting for Derivative Instruments and Hedging Activities.* Statement of Financial Accounting Standards No. 133. Norwalk, CT: FASB.

———. 1999. *Preliminary Views on Major Issues Related to Reporting Financial Instruments and Certain Related Assets and Liabilities at Fair Value.* Norwalk, CT: FASB.

Galai, D., and R. W. Masulis. 1976. The option pricing model and the risk factor of stock. *Journal of Financial Economics* (3): 53–81.

Gebhardt, W., C. M. C. Lee, and B. Swaminathan. 2001. Toward an implied cost of equity. *Journal of Accounting Research* (39): 135–176.

Gode, D., and P. Mohanram. 2003. Inferring the cost of equity using the Ohlson-Jüettner model. *Review of Accounting Studies* (8): 399–431.

Goh, J. C., and L. H. Ederington. 1993. Is a bond rating downgrade bad news, good news, or no news for stockholders? *Journal of Finance* (48): 2001–2008

Hail, L., and C. Leuz. 2006. International differences in the cost of equity capital: Do legal institutions and securities regulation matter? *Journal of Accounting Research* (44): 485–532.

Hand, J. R. M., P. J. Hughes, and S. E. Sefcik. 1990. Insubstance defeasances: Security price reactions and motivations. *Journal of Accounting and Economics* (13): 47–89.

———, R. W. Holthausen, and R. W. Leftwich. 1992. The effect of bond rating agency announcements on bond and stock prices. *The Journal of Finance* (47): 733–752.

Hayn, C. 1995. The information content of losses. *Journal of Accounting and Economics* (20): 125–153.

Hillegeist, S. A., E. K. Keating, D. P. Cram, and K. G. Lundstedt. 2004. Assessing the probability of bankruptcy. *Review of Accounting Studies* (9): 5–34.

Hodder, L., P. Hopkins, and J. Wahlen. 2006. Risk-relevance of fair value income measurement for commercial banks. *The Accounting Review* (81): 337–375.

Holthausen, R., and R. Leftwich. 1986. The effect of bond rating changes on common stock prices. *Journal of Financial Economics* (17): 57–90.

Huang, J. Z., and M. Huang. 2003. How much of the corporate-treasury yield spread is due to credit risk? Working paper, The Pennsylvania State University and Stanford University.

International Accounting Standards Board (IASB). 2003. *Financial Instruments: Recognition and Measurement.* IAS No. 39. London, U.K.: IASB.

———. 2006. *Preliminary Views on an improved Conceptual Framework for Financial Reporting—The Objective of Financial Reporting and Qualitative Characteristics of Decision-Useful Financial Reporting Information.* London, U.K.: IASB.

Jegadeesh, N. 1992. Does market risk really explain the size effect? *Journal of Financial and Quantitative Analysis* (27): 337–351.

Jorion, P., Z. Liu, and C. Shi. 2005. Informational effects of regulation FD: Evidence from rating agencies. *Journal of Financial Economics* (76): 309–330.

Kliger, D., and O. Sarig. 2000. The information value of bond ratings. *The Journal of Finance* (55): 2879–2902.

Kothari, S. P., and J. Zimmerman. 1995. Price and return models. *Journal of Accounting and Economics* (20): 155–192.

Lipe, R. 2002. Fair valuing debt turns deteriorating credit quality into positive signals for Boston Chicken. *Accounting Horizons* (16): 169–182.

Merton, R. C. 1974. On the pricing of corporate debt: The risk structure of interest rates. *The Journal of Finance* (29): 449–470.

Nelson, K. K. 1996. Fair value accounting for commercial banks: An empirical analysis of SFAS No. 107. *The Accounting Review* (71): 161–182.

Odders-White, E., and M. Ready. 2006. Credit ratings and stock liquidity. *Review of Financial Studies* (19): 119–157.

Parrino, R., and M. S. Weisbach. 1999. Measuring investment distortions arising from stockholder-bond holder conflicts. *Journal of Financial Economics* (53): 3–42.

Petrin, A., and K. Train. 2002. Omitted product attributes in discrete choice models. Working paper, University of Chicago and University of California.

Pinches, G., and J. C. Singleton. 1978. The adjustment of stock prices to bond rating changes. *The Journal of Finance* (33): 29–44.

Richardson, S., S. Teoh, and P. Wysocki. 2004. The walkdown to beatable analyst forecasts: The roles of equity issuance and insider trading incentives. *Contemporary Accounting Research* (21): 885–924.

Simko, P. 1999. Financial instrument fair values and nonfinancial firms. *Journal of Accounting, Auditing, and Finance* (14): 247–275.

Strong, J. S. 1990. Valuation effects of holding gains in long-term debt. *Journal of Accounting and Economics* (13): 267–283.

Vassalou, M., and Y. Xing. 2004. Default risk in equity returns. *The Journal of Finance* (49): 831–868.

White, H. 1980. A heteroscedasticity-consistent covariance matrix estimator and a direct test for heteroscedasticity. *Econometrica* (48): 817–838.

THE ACCOUNTING REVIEW
Vol. 83, No. 3
2008
pp. 665–703

Internal Control Weaknesses and Information Uncertainty

Messod Daniel Beneish
Indiana University

Mary Brooke Billings
New York University

Leslie D. Hodder
Indiana University

ABSTRACT: We analyze a sample of 330 firms making unaudited disclosures required by Section 302 and 383 firms making audited disclosures required by Section 404 of the Sarbanes-Oxley Act. We find that Section 302 disclosures are associated with negative announcement abnormal returns of −1.8 percent, and that firms experience an abnormal increase in equity cost of capital of 68 basis points. We conclude that Section 302 disclosures are informative and point to lower credibility of disclosing firms' financial reporting. In contrast, we find that Section 404 disclosures have no noticeable impact on stock prices or firms' cost of capital. Further, we find that auditor quality attenuates the negative response to Section 302 disclosures and that accelerated filers—larger firms required to file under Section 404—have significantly less negative returns (−1.10 percent) than non-accelerated filers (−4.22 percent). The findings have implications for the debate about whether to implement a scaled securities regulation system for smaller public companies: material weakness disclosures are more informative for smaller firms that likely have higher pre-disclosure information uncertainty.

Keywords: *information uncertainty; reporting credibility; Sarbanes-Oxley; scaled securities regulation.*

Data Availability: *Data are available from public sources.*

JEL Classification: *G14; G18; G30.*

We thank the editor and two anonymous reviewers for their help in improving the paper. We benefited from many useful comments from participants at the June 2005 Summer Conference Series at Purdue University, the 2005 Conference on Financial Economics and Accounting (particularly discussant Jere Francis), and the 2005 SEC/AICPA Conference on Financial Reporting. We thank Jennifer Francis for supplying the factor-mimicking portfolio data used to calculate e-loadings and Sarah McVay for providing coding for weakness severity. We also thank I/B/E/S International, Inc. for providing forecast data used in this study.

Editor's note: This paper was accepted by Dan Dhaliwal.

Submitted April 2006
Accepted October 2007

I. INTRODUCTION

A series of Congressional inquiries into the integrity of financial reporting in the United States culminated with the passage of the Sarbanes-Oxley Act (SOX or Act) on July 30, 2002. The degree to which SOX transforms the rules of corporate governance, disclosure, and financial reporting, and the high compliance costs it imposes on SEC registrants cause many to question the benefits of this extensive regulatory intervention (Ribstein 2002; Michaels 2003; Romano 2005; DeFond and Francis 2005).

A specific objective of Sections 302 and 404 of SOX is to inform investors about weaknesses in disclosing firms' systems of internal controls that may increase financial statement errors or managers' ability to manage earnings. If investors already recognize that disclosing firms have poor earnings quality and discount stock prices as compensation for greater uncertainty, then disclosures about internal control weaknesses are unlikely to be incrementally informative. Alternatively, if weakness disclosures signal incremental risk, or if investors do not fully impound observable, value-relevant information into stock prices, then weakness disclosures are likely to be associated with adverse stock market responses as investors revise upward their risk assessments. An adverse response is consistent with recent studies that suggest that investors require compensation for uncertainty about firms' financial reports (Easley and O'Hara 2004; Lambert et al. 2007a; Ecker et al. 2006).

A number of contemporaneous studies use disclosures under Sections 302 and 404 of SOX to investigate firms that disclose internal control weaknesses. These studies show that, *ex ante*, firms disclosing internal control weaknesses are smaller, younger, riskier, and financially weaker (Ashbaugh-Skaife, Collins, and Kinney 2007; Doyle et al. 2007a), have poorer *ex ante* accrual and earnings quality (Ashbaugh-Skaife et al. 2008; Bedard 2006; Doyle et al. 2007b) and higher costs of equity capital (Ashbaugh-Skaife, Collins, Kinney, and LaFond 2007; Ogneva et al. 2007). What these studies do not show is whether the disclosures of material weaknesses under 302 and 404 convey information to investors, and, in the affirmative, the nature of the information they convey. This paper addresses these questions.

Specifically, we analyze the market reactions to Sections 302 and 404 disclosures both unconditionally and conditional on prior disclosure and the richness of the firm's information environment. We also investigate whether forecasted earnings per share and implied costs of capital change after a material weakness disclosure. We base our analyses on a sample of 330 firms making unaudited material weakness disclosures under Section 302, and 383 firms disclosing material weaknesses in audited Section 404 reports. We benchmark our return and costs of capital analyses against specific control samples of firms that have similar characteristics to weakness disclosers (i.e., small, financially weaker, poorer earnings quality) but that do not disclose internal control weaknesses.

Our returns analyses reveal the following. First, we find a significant adverse impact to disclosures of unaudited material weaknesses under Section 302: the three-day abnormal return equals −1.8 percent and tests of subsequent window abnormal returns suggest that the negative market impact is not transitory. Section 302's impact is consistent with investors requiring compensation for holding firms with poorer financial reporting quality. Second, we detect no market response to Section 404 disclosures, both unconditionally and conditional on prior Section 302 disclosures. Our inability to detect a market response to Section 404 disclosures is consistent with the hypothesis that accelerated filers (the larger firms that are required to file under Section 404) operate in richer information environments, as well as Doyle et al.'s (2007b) hypothesis that audited internal control disclosures reflect a lower materiality threshold for disclosure. However, we cannot completely eliminate the possibility that data constraints reduce the power of our tests to detect a response.

Third, our analysis of Section 302 disclosures conditional on Section 404 filing status reveals that accelerated filers experience significantly less negative returns (−1.10 percent) than non-accelerated filers (−4.22 percent). This suggests that Section 302 disclosures are more informative for firms that likely have higher pre-disclosure information uncertainty. This result has implications for the debate about whether a scaled securities regulation system for smaller public companies should be implemented (SEC Advisory Committee on Smaller Public Companies 2006): had Section 302 exemptions mirrored Section 404 exemptions in terms of firm size, the material weakness disclosures would, at best, be weakly informative.

Fourth, in a cross-sectional analysis of the market response to Section 302 disclosures, we find no association between market response and disclosure source (10-K, 8-K, 10-Q, or other) or the severity of the weakness. However, we do find that auditor quality attenuates the negative market response to Section 302 disclosures. This finding suggests that a shock to financial reporting credibility is attenuated by the presence of a high-quality auditor, even when the disclosures themselves are unaudited.

Finally, because the adverse market reaction to Section 302 disclosures can reflect an increase in risk, a decrease in expected future cash flows, or both, we collect earnings forecast revisions to determine whether earnings expectations decline for disclosing firms, and we use observed price changes and forecast revisions to estimate changes in implied cost of capital with four models commonly used in prior work (Dhaliwal et al. 2006; Ogneva et al. 2007). We find that Section 302 disclosers experience abnormally negative forecast revisions and abnormally positive increases in costs of capital after their Section 302 disclosure. We do not find these results for firms making disclosures under Section 404. The results for Section 302 suggest that disclosing firms experience negative changes in expected cash flows as well as increases in perceived risk.

Interpreting our results requires cautions about our ability to correctly identify disclosure dates and to control for contaminating information released contemporaneously, and about our reliance on analysts' expectations of future earnings in solving for firms' implied costs of capital. Subject to these cautions, the evidence suggests that SOX 302 disclosures are informative and point to lower credibility of disclosing firms' financial reporting. However, we find no evidence that SOX 404 disclosures are informative. Our inability to detect information content or changes in costs of capital surrounding Section 404 disclosures is consistent with evidence in Doyle et al. (2007b) showing no relation between 404 internal control weaknesses and accrual quality and evidence in Ogneva et al. (2007) showing no difference in cost of capital levels between firms that disclose weaknesses under 404 and those that do not.

As discussed more fully in Section II, our study adds to contemporaneous research on the effects of SOX by quantifying the market reaction to weakness disclosures and measuring changes in forecasts and implied costs of capital. As well, our analyses of both Section 302 and Section 404 disclosures separately and jointly add to concurrent work that focuses on Section 302 weakness disclosures, either to investigate whether these disclosures benefit small or large investors (De Franco et al. 2005), or to determine whether the information content of the disclosures increases with the severity of the control weakness (Hammersley et al. 2008).

Our study also provides evidence linking financial reporting uncertainty to stock prices. Prior archival research draws inferences about reporting quality from firms' routine financial disclosures and results of operations (e.g., Botosan 1997; Ecker et al. 2006). In contrast, our setting allows us to evaluate the capital market response to disclosures that specifically address the integrity of individual firms' financial reporting processes. As well, by examining the marginal effect of auditor association with the firm on the stock market response to unaudited disclosures, our study adds to research investigating how auditor quality affects the perceived credibility of firms' financial reports (Teoh and Wong 1993; Becker et al. 1998).

The remainder of the paper progresses as follows. Section II of the paper provides background on the Act and discusses existing literature examining information risk. Section III supplies our hypotheses and research design. Section IV describes the sample selection and data collection and provides descriptive statistics. Section V presents the analyses and results of the study and Section VI concludes.

II. BACKGROUND AND RELATED RESEARCH

Regulation of Internal Accounting Controls

Prior to the passage of SOX, the SEC had little authority over firms' operational matters or their internal accounting controls. As we summarize in Figure 1, SOX significantly changes the information environment of public companies by requiring disclosures about firms' internal controls under Sections 302 and 404. First effective in 2002, Section 302 requires managers (1) to certify that weaknesses and deficiencies have been privately reported to the audit committee and (2) to disclose material weaknesses and material changes in internal control to the public.[1] Because Section 302 does not require attestation by an independent auditor and, instead, only requires disclosure of material weaknesses about which the CEO or CFO has knowledge, disclosure under Section 302 is a function of managers' diligence in identifying weaknesses and their discretion over disclosure.

In contrast, Section 404 imposes affirmative responsibilities on management and the company's independent auditor to detect material weaknesses in internal controls. Specifically, Section 404 requires management to document and test the efficacy of internal controls and to issue annually an internal control report in which management must make either positive assertions concerning the effectiveness of internal controls, or disclose the nature of deficiencies and weaknesses that render controls ineffective. The company's auditor must issue a separate opinion over management's assertions to be included in the annual report. Section 404 is effective for accelerated filers for fiscal years ending after November 15, 2004.[2]

Related Research

The passage of SOX constitutes a structural shift in the information environment, around which research is burgeoning. Early research documents a positive market reaction to the regulatory proposals that culminated with the passage of the Act, consistent with the legislative action having a net beneficial effect on the perceived quality of financial reporting (Jain and Rezaee 2006; Jain et al. 2008; Li et al. 2008). In contrast to studies that evaluate the effect on the aggregate information environment, we assess the implications for individual firms of the implementation of a specific provision of the Act—the requirement to disclose known material weaknesses over internal controls.

Figure 2 presents a schematic of the environment shift in event-time, together with selected research studies most relevant to our study. The first column in Figure 2 shows

---

[1] As discussed in further detail in Section IV, firms disclose Section 302 weaknesses on a variety of SEC forms, including annual and quarterly reports, forms 8-K, registration statements, and proxy statements.

[2] Accelerated filers are generally those domestic firms with a market capitalization of $75 million or greater.

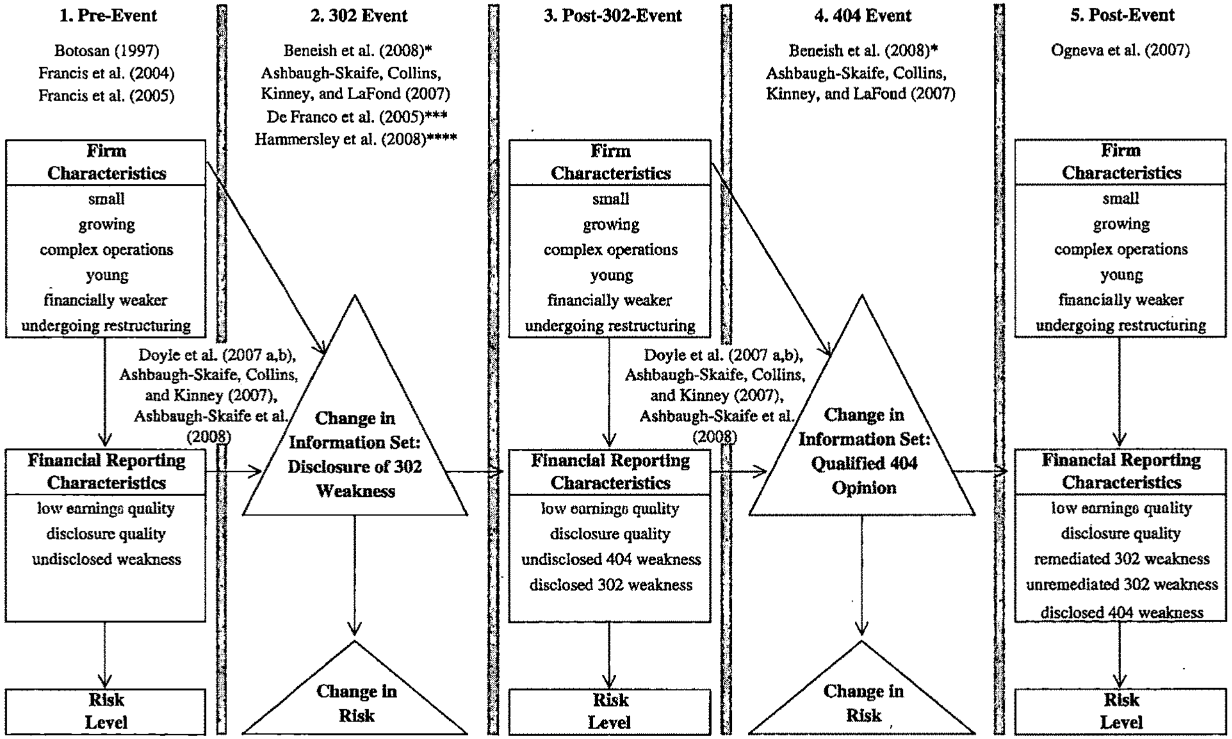
FIGURE 1
Summary of Selected Sarbanes-Oxley Provisions

The Sarbanes-Oxley Act ("SOX")	July 30, 2002. President Bush signs SOX into law.	SOX revolutionizes the rules of corporate governance, disclosure, and financial reporting for all SEC registrants.
Sections 101 & 103—*Public Company Accounting Oversight Board and Auditing Standards*	Section 101 of SOX creates the Public Company Accounting Oversight Board (PCAOB).	Section 103 charges the PCAOB with establishing auditing and related attestation, quality control, ethics, and independence standards and rules to be used by registered public accounting firms in the preparation and issuance of the audit report.
Section 906—*White collar crime penalty, corporate responsibility for financial reports*	Section 906, effective immediately upon passage, is the one of three sections of SOX requiring certification.	Under 906, management must certify that 10-Ks and 10-Qs are fairly stated and are in compliance with the Securities and Exchange Act of 1934. Knowing and/or willful wrongdoing is subject to fines and imprisonment.
Section 302—*Corporate responsibility for financial reports*	August 29, 2002. Issued by the SEC on August 27, 2002, the rule becomes effective on August 29, 2002.	Executives must certify that: • they reviewed the report, • the report contains no misrepresentations, • the financial information is fairly presented, • they have reported any internal control weaknesses (including fraud) to the audit committee, and • they have identified material changes in internal controls.
Auditing Standard No. 2—*An Audit of Internal Control over Financial Reporting Performed in Conjunction with an Audit of Financial Statements*	June 17, 2004. Issued by the PCAOB on March 9, 2004, the SEC approves Audit Standard No. 2 on June 17, 2004.	Auditing Standard No. 2, effective for audits of internal control over financial reporting required by Section 404(b), requires that the auditor's report on internal control over financial reporting include the auditor's opinion on management's assessment of the effectiveness of internal control over financial reporting as well as the auditor's independent assessment of the effectiveness of the company's internal control over financial reporting.
Section 404—*Management assessment of internal controls*	November 15, 2004. The SEC implements the rule on May 27, 2003, and extends the compliance dates on February 24, 2004.	Under 404(a), management of the company must supply an "internal control report" that that takes responsibility for adequate internal controls and assesses the effectiveness of the controls over financial reporting. Under 404(b), the company's auditor must attest to, and report on, management's assessment of internal controls. Companies have 90 days from the end of their fiscal year to comply. November 15, 2004, marks the start for companies with market capitalization greater than $75 million. July 15, 2005, marks the start for companies with less than $75 million market capitalization.

FIGURE 2

Research Examining Weakness Disclosures under Sections 302 and 404 of the Sarbanes-Oxley Act of 2002

Do weakness disclosures convey incremental information about financial reporting quality to investors?



---

* Beneish et al. (2008) calculate size-adjusted cumulative abnormal returns during the three-day window surrounding weakness disclosures and then compare changes in four estimates of cost of capital (following Dhaliwal et al. 2006) for weakness firms to those of control firms matched on industry and earnings quality.

** Ashbaugh-Skaife, Collins, Kinney, and LaFond (2007) compare changes in estimated rates of return (reported by Value Line) for material weakness firms to those of the remaining population of Value Line firms.

*** De Franco et al. (2005) evaluate wealth redistribution effects of weakness disclosure among classes of shareholders.

**** Hammersley et al. (2008) examine the market reaction to weakness disclosures and their associated characteristics for firms disclosing weaknesses under Section 302.

that prior to the passage of SOX, firm characteristics and reporting quality are associated with risk levels.[3] For example, Botosan (1997) relates higher disclosure quality to lower costs of capital and Francis et al. (2004, 2005) relate poorer earnings quality with higher costs of capital. Consistent with this research, the last column shows that firm and reporting characteristics, including earnings quality, are related to cost of capital levels subsequent to SOX (Ogneva et al. 2007; Ashbaugh-Skaife, Collins, Kinney, and LaFond 2007). However, results are mixed as to whether prior disclosure of a material weakness (weakness status) provides incremental explanatory power for cost of capital levels. Ogneva et al. (2007) find that weakness status derived from Section 404 disclosures has little or no explanatory power for cost of capital levels, after controlling for firm characteristics and other components of reporting quality. In contrast, Ashbaugh-Skaife, Collins, Kinney, and LaFond (2007) find a positive relation between cost of capital levels and weakness status derived from Section 302 and Section 404 disclosures. This inconsistency suggests that additional research is required to understand the relation between weakness disclosures and perceived risk.

An association between cost of capital *levels* and weakness status does not necessarily imply that the market responds to the disclosure of internal control weaknesses. Indeed, Ge and McVay (2005), Doyle et al. (2007a, 2007b), Ashbaugh-Skaife, Collins, and Kinney (2007), and Ashbaugh-Skaife, Collins, Kinney, and LaFond (2007) find that firms disclosing weaknesses have poor reporting quality prior to disclosure, and that such weaknesses are associated with externally observable firm characteristics. If stock prices are fully adjusted for poor reporting quality based on observable traits, then weakness disclosures convey no new information to the market. Thus, the information effects of weakness disclosures are unclear.

To assess the effects of weakness disclosures on risk, we evaluate the market reaction to reports made under Sections 302 and 404. We represent these reporting events in the second and fourth columns of Figure 2, respectively. We focus on *changes* in risk (rather than risk levels) because doing so helps to rule out alternative explanations for observed differences between disclosers and nondisclosers, as firm attributes (and other indicators of reporting quality) likely remain unchanged over shorter intervals. However, this design decision requires us to make assumptions about the duration of the risk change and the timeliness of analysts' forecasts.

Several contemporaneous working papers also examine the effects of firm-specific risk disclosures on the market. De Franco et al. (2005) evaluate wealth redistribution effects of weakness disclosure among classes of shareholders (small and large investors). Relying on TAQ data, De Franco et al. (2005) show that net selling by small investors drives the market's negative response to weakness disclosures and conclude that small investors benefit more from material weakness disclosures than large investors.[4]

Shifting focus to the characteristics of the 302 weakness disclosures, Hammersley et al. (2008) suggest that the information content of Section 302 disclosures depends on weakness severity: they detect no reaction to disclosures of control deficiencies, and increasingly negative reactions to disclosures of significant deficiencies and material weaknesses.[5] In related work, Ashbaugh-Skaife, Collins, Kinney, and LaFond (2007) examine changes in Value Line's Projected 3- to 5-Year Average Return (PAR) incident to weakness disclosures and firms' subsequent remediation efforts and find that internal control risk affects the market's assessment of firms' costs of capital.

---

[3] Firm characteristics include attributes such as size, age, growth rate, operational complexity, financial strength, and restructuring. We use the term "reporting quality" to mean completeness of disclosure (e.g., Botosan 1997) as well as earnings quality (e.g., Francis et al. 2005) and information about the adequacy of internal controls over reporting.

[4] Specifically, De Franco et al. (2005) examines a sample of 102 firms reporting internal control weaknesses (which include 45 uncontaminated material weaknesses disclosures) and document negative abnormal returns to material weakness disclosures.

Our study differs from contemporaneous studies in several notable respects. First, we measure changes in analysts' forecasts and the implied costs of capital in an attempt to separately measure possible cash flow implications and risk changes. Moreover, because disclosing firms have observably poor earnings quality prior to disclosure, we compare the returns of disclosing firms to those of a control sample of nondisclosing firms with similarly poor earnings quality.[6] We also assess the persistence of the market's response by evaluating abnormal returns over longer windows subsequent to the 302 disclosure. In doing so, we separately evaluate the market response to firms' subsequent Section 404 disclosures, and ascertain that Section 302 and Section 404 disclosures are not informationally equivalent. Finally, we examine the effects of disclosure context, auditor quality, and an alternative measure of weakness severity developed by Doyle et al. (2007b) on the market's response to weakness disclosures.

Although Ashbaugh-Skaife, Collins, Kinney, and LaFond (2007) also measure changes in required returns, our sample and methodology for measuring cost of capital differ, and certain of our results are inconsistent with, or qualify, their findings. We detect no evidence of abnormal returns or changes in implied cost of capital incident to firms' Section 404 disclosures, including disclosures that suggest remediation of previously disclosed weaknesses. We attribute some of these inconsistencies to our use of a matched control sample for computing abnormal returns and changes in cost of capital, and our use of implied cost of capital rather than PAR.[7] Consistent with Ogneva et al. (2007), we note that the use of PAR results in sample reduction and potentially biases results.[8]

Our study also contributes to research linking auditor quality and the perceived credibility of financial reports (e.g., Teoh and Wong 1993; Becker et al. 1998). Although Section 302 disclosures are unaudited, firms making disclosures are required to engage an auditor for their financial statements, and accelerated filers must solicit a separate opinion over internal controls. Engagement of high-quality auditors may mitigate the market response to discretionary Section 302 weakness disclosures if their interim auditing programs impose on their clients a lower threshold for disclosure. Ge and McVay (2005) find that firms engaging high-quality auditors are more likely to make disclosures of material internal control weaknesses. Ge and McVay (2005) leave open the question of whether large accounting firms are associated with material weakness disclosures because companies more diligent about disclosure tend to engage high-quality auditors, or whether high-quality auditors tend to take on clients with greater reporting risk. We provide insight into this question by assessing the effect of large firm engagement on the market response to unaudited material weakness disclosures. Association with high-quality auditors will mitigate the market response to disclosure if large auditor clients have a lower threshold for disclosure, but not if large auditor clients are simply more risky.

---

[5] Hammersley et al. (2008) also report that the stock price reaction is more adverse when management concludes that internal controls are effective despite the existence of an internal control weakness, and when internal control weaknesses are less auditable.

[6] The control sample provides us a measure of abnormal return relative to a portfolio of size-, industry-, and earnings-quality matched firms. Our inferences are robust to a matched-pair comparison based on industry, size, and earnings quality.

[7] In an attempt to replicate Ashbaugh-Skaife, Collins, Kinney, and LaFond's (2007) findings, we collect PAR for our sample of initial material weakness disclosers. This results in a 52 percent reduction in sample size. We find that PAR increases for this sample of firms relative to all others for which PAR is reported; however, increases in PAR for disclosing firms do not differ significantly from an industry- and earnings-quality-matched control sample. This suggests that inferences based on PAR depend on the benchmark used to market-adjust changes in cost of capital.

[8] As an expected annualized price change over a finite horizon, PAR may be negative for firms that analysts believe are overpriced relative to projected earnings. It is not clear to us how these observations should be treated, so we delete them. On average, Value Line firms are larger, more profitable, have greater analyst following and are more likely to be audited by a Big 4 auditor. A control sample comprised of this population is not likely to represent firms with more opaque information environments.

III. HYPOTHESES AND RESEARCH DESIGN

The Capital Market Consequences of Material Weakness Disclosures

One intention of Section 302 is to inform investors of weaknesses in disclosing firms' systems of internal controls. If weak control mechanisms increase measurement errors or managers' ability to manage earnings, then these disclosures may identify firms with previously unknown low quality financial reporting. Theoretically and empirically, recent research indicates that investors require compensation for uncertainty about a firm's financial reporting quality. These studies argue that uninformed investors face an adverse selection problem in trading securities in the capital market (Easley and O'Hara 2004; Francis et al. 2005; Ecker et al. 2006; Lambert et al. 2007a, 2007b).[9]

If investors require compensation for uncertainty about a firm's financial reporting quality, then weakness disclosures have the potential to cause investors to revise upward their beliefs about firm risk. Therefore, we expect disclosing firms to experience both negative abnormal returns and an increase in their cost of capital in response to lower perceived reporting quality. However, investor belief revision requires either that the disclosures convey new information to the market or that the disclosures cause investors to place more weight on value-relevant information previously overlooked in valuation (e.g., Ou and Penman 1989; Sloan 1996).

There is some question as to whether internal control weakness disclosures meet the necessary conditions to cause belief revision. If weakness disclosures do not convey any incremental information about the quality of firms' reporting or future cash flows, and if the risk and cash flow expectations associated with low-quality reporting are fully impounded in price, then the public disclosure of weaknesses is unlikely to have informational value. Our first hypothesis, stated in the alternative, reflects this empirical question:

> **H1:** Material weakness disclosures cause investor belief revision that is reflected in cumulative abnormal returns.

We evaluate the stock market reaction to Section 302 weakness disclosures by conducting an event study over the three-day window surrounding the disclosure date for a sample of 330 firms disclosing weaknesses under Section 302 and three control samples matched on industry and either size (total assets), performance (return on assets), or *ELOAD* (the market-based earnings quality measure introduced by Ecker et al. 2006).[10] We perform a similar analysis for a sample of 383 firms disclosing control weaknesses under Section 404. Material weaknesses are frequently disclosed contemporaneously with other value-relevant information. Consequently, we perform a detailed contamination analysis for both disclosing and control firms.[11] Controlling for potential confounds increases the likelihood that the observed reaction is due to the material weakness disclosure. Further, because we posit that Section 302 and 404 disclosures are related events, we structure tests of the market reaction to (1) Section 302 disclosures conditional on firms' requirement to subsequently file Section 404 reports, and (2) Section 404 disclosures conditional on firms' prior Section 302 disclosures.

---

[9] For example, Francis et al. (2005) and Ecker et al. (2006) find that a proxy for earnings quality explains variation in firms' costs of debt and equity, and show that including a factor based on such a proxy for earnings quality explains variation in future stock returns incrementally to the three factors proposed by Fama and French (1992). Other studies argue that information availability and information precision determinants of equilibrium prices. For example, Arbel and Strebel (1982), Barry and Brown (1984), and Merton (1987) suggest that investors demand higher returns for holding stocks with less available information. Beneish and Gardner (1995) show that the market reaction experienced by firms deleted from the Dow Jones Industrial Average is consistent with the investors demanding a premium for higher trading costs and for holding securities that have relatively less available information. Beneish et al. (2005) describe conditions under which auditor resignations announcements revise investors' beliefs about the quality of a firms' financial reporting in excess of the credibility discount already embedded in price.

The market reaction to Section 302 disclosures may vary with firms' Section 404 accelerated filing status for two reasons. First, Doyle et al. (2007b) suggest that Section 404's auditor attestation requirement results in a lower threshold for disclosure under Section 404 than Section 302. Because accelerated filers are aware that the internal controls will be subsequently audited and any weaknesses disclosed under Section 404, accelerated filers may also have a lower threshold for disclosure when making unaudited disclosures under Section 302. Second, the market response may be more extreme for non-accelerated filers because of greater uncertainty in their information environments. This is because, by statutory construction, non-accelerated filers are smaller, and such firms are presumed to have less sophisticated information systems and less analyst following.[12]

The market response to Section 404 disclosures also may vary conditionally with firms' prior Section 302 disclosures. We hypothesize that the market response will be greatest for Section 404 disclosures that disconfirm prior Section 302 disclosures. Specifically, "effective" Section 404 disclosures may be interpreted as positive events for firms previously reporting weaknesses under Section 302 (Ashbaugh-Skaife, Collins, Kinney, and LaFond 2007), but may provide little incremental information otherwise. Similarly, "ineffective" Section 404 disclosures may be interpreted as negative surprises for firms previously not reporting such weaknesses, but may provide little incremental information when firms have previously reported the weaknesses. Structuring our analyses to incorporate conditional responses should increase the power of our tests and allow us to qualify contemporaneous research, some of which either pools Section 404 and Section 302 disclosures, or draws inferences about material weakness disclosures from Section 404 alone.

An adverse return reaction can reflect an increase in the expected rate of return and/ or lower expected future cash flows. Reduced expected future cash flows can be related to costs directly associated with the material weakness (e.g., expected costs of remediating the control weakness, expected increased audit fees, etc.). Costs can also be indirect, if the material weakness leads customers, suppliers, and other stakeholders to question the firm's ability to fulfill its implied commitments (Bowen et al. 1995; Matsumoto 2002). To partition the return reaction into cash flow and risk components, we conduct several additional analyses. First, we identify abnormal changes in expected cash flows by comparing earnings forecast revisions for disclosing firms to those of various control samples comprised of firms matched on industry and each of earnings quality, size, or performance. If disclosing firms are expected to experience higher direct and indirect costs, we should observe less positive forecast revisions for these firms. Although we cannot isolate changes in indirect cost components, we can use the same comparisons to identify certain abnormal increases in direct costs for disclosing firms, such as audit and professional fees. As reflected in Hypothesis 2, we posit that disclosing firms experience lower expected future cash flows and higher increases in audit and professional fees than comparable nondisclosing firms.

---

[10] We compute size-adjusted cumulative abnormal returns over the three-day window surrounding the material weakness disclosure by subtracting from individual firm daily returns the concurrent equally weighted market return for other firms in the same market capitalization decile.

[11] We discuss the selection of event dates and the nature of contaminating information in Section IV.

[12] In addition, in contrast to Section 302 disclosures, Section 404 disclosures generally accompany annual reports and exhibit greater event-date clustering, which may introduce noise into the event-period returns and, consequently, limit our ability to detect a reaction.

> **H2:** Firms disclosing internal control weaknesses experience less positive forecast revisions and higher audit and professional fees than comparable firms.

We use analysts' forecasts before and after the material weakness disclosure to estimate changes in implied costs of capital following the approach suggested by Hribar and Jenkins (2004). A change in this measure suggests a change in investors' risk perceptions. However, this interpretation is subject to the timeliness of analysts' forecasts as a proxy for earnings expectations. Guay et al. (2006) posit that systematic errors in implied costs of capital are induced by analyst forecasts that do not incorporate in a timely manner changes in earnings expectations reflected in returns. We cannot eliminate the possibility that potentially untimely forecasts are inducing error in our measure of cost of capital changes.[13] However, we perform a series of robustness tests to evaluate the effect of potential measurement error on our findings.

To compute cash flow effects, we use a matched control sample that permits a measure of abnormal forecast revisions and abnormal changes in cost of capital. As summarized in Hypothesis 3, we predict that as well as having lower expected future cash flows, the perceived risk of disclosing firms will increase by more than that of comparable firms.

> **H3:** Firms disclosing control weaknesses experience higher increases in implied costs of capital than comparable firms.

As described in the Appendix, we estimate equity cost of capital (*ECC*) based on residual income models that specify a relation between equity cost of capital, equity market value, equity book value, and forecasted earnings and dividends. To mitigate the effects of error associated with estimating *ECC*, we eliminate observations for which $ECC < 0$ percent or $ECC > 50$ percent.[14]

---

[13] Guay et al. (2006) show that sluggish forecast revisions result in downward (upward) bias following large positive (negative) stock returns over the year preceding measurement. We measure price change as the difference between the average price over the five days subsequent to the event window and the average price over the five days prior to the event window. If forecast revisions fail to reflect decreased earnings expectations embedded in event window price declines, then the implied cost of capital is likely to be biased upward. For this reason, we measure analyst forecast revisions over various periods up to 180 days before and after the observed price change to allow additional time for forecasts to be revised. As well, we confirm the robustness of our findings to weighting costs of capital by *ex post* forecast accuracy. If forecast accuracy is negatively correlated with measurement error, then this should mitigate the effects of measurement error.

[14] In his discussion of Hribar and Jenkins (2004), Kasznik (2004) highlights a number of methodological concerns and offers suggestions for future research. Following Kasznik (2004), we use the detail I/B/E/S database to supply analysts' forecasts of earnings per share in year one and year two, rather than the summary file.

Cross-Sectional Variation in Section 302 Abnormal Returns

Because Section 302 disclosures reflect discretion with respect to timing and content, we use the full sample to investigate whether the hypothesized market reaction varies according to auditor quality (*AQ*) and other factors that may influence investor responses to disclosure. Inclusion of *AQ* follows prior research that investigates how auditor quality affects the perceived credibility of financial reports. We expect that high auditor quality mitigates negative effects of increased information uncertainty, either because auditors improve the precision and credibility of financial disclosures (Francis et al. 2004; Palmrose et al. 2004; Teoh and Wong 1993) or because well-capitalized audit firms provide potential indemnification against losses suffered due to potential misstatements (Bockus and Gigler 1998).

Consistent with this literature, we set audit quality (*BIG*) equal to 1 when the audit firm is a member of the Big 4, and 0 otherwise. In addition to this simplified measure of audit quality, we calculate measures of auditor industry specialization based on each accounting firm's industry market share (Craswell and Taylor 1991; Craswell et al. 1995). Specifically, we code an indicator variable (*INDSPEC*) equal to 1 if the firm is audited by an industry specialist, as measured by a threshold of 20 percent of market share in industries with at least 30 firms. If the firm engages a Big 4 auditor that is not an industry specialist, we set an indicator variable (*NONSPEC_BIG*) equal to 1. We examine the marginal effect of auditor type on the market reaction associated with material weakness disclosure using the following regression:

$$\begin{aligned} CAR_{i,t} = {} & \alpha_0 + \alpha_1 AQ_{i,t} + \alpha_2 RESIGN_{i,t} + \alpha_3 AUDCHGE_{i,t} \\ & + \alpha_4 CONTAMINATION_{i,t} + \alpha_5 CONTEXT_{i,t} + \alpha_6 NO404REQ \\ & + \alpha_7 CLW + \alpha_8 ELOAD_{i,t} + \alpha_9 FEECHGE_{i,t} + \varepsilon_{i,t}. \end{aligned} \tag{1}$$

We predict a positive coefficient for *AQ*, with either *BIG* or *NONSPEC_BIG* and *INDSPEC* as alternative proxies for *AQ*. Because we expect the market response to vary according to firms' Section 404 filing status, we code an indicator variable (*NO404REQ*) equal to 1 if the firm is a non-accelerated filer. Consistent with non-accelerated filers operating in poorer information environments, we predict a negative coefficient for *NO404REQ*.

The severity of the disclosed weakness may influence the market response as well. Hammersley et al. (2008) predict that disclosure of more severe weaknesses will result in more negative investor responses. However, their prediction requires disclosures to be equally unanticipated by the market, independent of the severity of the weakness. Doyle et al. (2007a, 2007b) find that firms disclosing more severe "company-level" weaknesses exhibit distinctive characteristics that are observable prior to disclosure. Relying on classifications supplied by Doyle et al. (2007a, 2007b), we set an indicator variable (*CLW*) equal to 1 if the firm disclosed a company-level weakness and 0 otherwise. We do not predict the sign of the coefficient on *CLW*, because in contrast to Hammersley et al. (2008), Doyle et al. (2007b) and Ashbaugh-Skaife, Collins, Kinney, and LaFond (2007) posit that the changes in risk resulting from negative signals about firms' reporting quality will be less severe for weaknesses that were more visible *ex ante*.

We control for pre-disclosure differences in reporting quality with the variable *ELOAD*. Ecker et al. (2006) identify an earnings quality factor as one important determinant of returns, and provides evidence that *ELOAD* captures a firm's sensitivity to poor earnings quality.[15] If investors distinguish firms' pre-disclosure earnings quality and discount stock prices according to estimates of *ELOAD*, then we expect weakness disclosures to be less informative for poor earnings quality firms (high *ELOAD*). Therefore, we expect a positive coefficient on the *ELOAD* variable.

We include post-disclosure changes in audit and professional fees to determine whether the expected differential fee impact across firms explains some of the cross sectional variation in announcement abnormal returns. The variable *FEECHGE* is computed as the percentage change in audit fees or total professional fees between 2003 and 2004, adjusted by median fee changes for all firms with available data.[16] We expect a negative coefficient on *FEECHGE* to the extent that this control variable serves as a proxy for the higher costs directly associated with the material weakness and these are reflected in announcement returns.

We include the variables *CONTEXT*, *RESIGN*, *AUDCHGE*, and *CONTAMINATION* to represent information that is released contemporaneously to material weakness disclosures under Section 302. *CONTEXT* comprises three indicator variables representing the source of disclosure.[17] We do not predict a sign for the *CONTEXT* variables because we include controls for contaminating information frequently associated with *CONTEXT*. *RESIGN* is a variable indicating that the current auditor has resigned, and *AUDCHGE* is a variable indicating that a new auditor has been appointed.[18] Relying on prior auditor resignation research, we predict a negative coefficient for *RESIGN*. We make no prediction for *AUDCHGE*. Similarly, we make no prediction for *CONTAMINATION*, which indicates a variety of other events contemporaneously announced (e.g., earnings announcements, litigation updates, officer changes, restatement updates, bond rating changes) for which expectations are difficult to specify.

IV. DATA COLLECTION AND DESCRIPTIVE STATISTICS

We compile a sample of 330 firms disclosing material weaknesses in internal control under Section 302 of SOX, which was effective on August 29, 2002. We perform a full-text search of SEC filings for calendar year 2004 and collect text paragraphs containing the phrases "material weakness" and "internal control." After culling repeated disclosures and filings containing only generic references to Sarbanes-Oxley reporting requirements,

---

[15] We estimate *ELOAD* over the calendar year 2003 for the CRSP population running the following regression:

$$(R_{i,t} - R_t^f) = \alpha_i + \beta_i(R_{M,t} - R_t^f) + s_i SMB_t + h_i HML_t + ELOAD_i EQF_t + \varepsilon_{i,t}$$

where $R_{i,t}$ denotes the return to decile i for month t, R_t^f denotes the return on the 30-day T-bill for month t, $R_{M,t}$ denotes the return on the value-weighted CRSP index for month t, SMB_t denotes the Fama-French size factor mimicking portfolio, HML_t denotes the Fama-French book-to-market factor mimicking portfolio, and EQF_t denotes the Ecker et al. (2006) earnings quality factor mimicking portfolio.

[16] To allow for the fact that auditor changes influence fees, we compute medians for several different groups ranging from no auditor change, as well as changes from Big 4 to non-Big 4, non-Big 4 to Big 4 and so on.

[17] Material weaknesses under Section 302 can be disclosed within a variety of different SEC filings depending on the timing of detection relative to periodic filing requirements. *CONTEXT* comprises three indicator variables. We set *CURR* equal to 1 if the disclosure is made on a current report (8-K), and 0 otherwise; *REG* is set equal to 1 if the disclosure is made on an a 1933 Act registration statement, and 0 otherwise; and *K* is set equal to 1 if the disclosure is made on an annual report, and 0 otherwise. Quarterly reports comprise the reference group.

[18] For example, if a sample firm announces an auditor resignation in the event window, then *RESIGN* is set to 1. If a firm announces the replacement of its current auditor with a new auditor, then *AUDCHGE* is set to 1.

523 unique firms making initial material weakness disclosures remain. Of those 523 firms, 187 firms either do not trade or trade on pink sheets.[19]

Panel A of Table 1 describes the Section 302 and 404 firms with available stock return data by month of filing and form type. Consistent with the second and third quarterly report (Form 10-Q) filing deadlines for calendar year firms, 12 (26) percent of the Section 302 material weakness disclosures occur in August (November). Similarly, approximately 72 percent of Section 404 reports are filed in March and April, reflecting the annual report due dates for firms with fiscal year ending in December and January. The grouping of disclosure dates suggests the need to adjust for clustering in the returns analyses. Untabulated results show that the majority of Section 302 disclosures are made in conjunction with Form 10-Q filings (158) and annual reports (78). All of the Section 404 reports are disclosed with the annual report.

We obtain stock return data and market capitalization from CRSP. Panel B of Table 1 shows that of the 330 Section 302 firms with available returns data, 59 percent trade on the National Market (NASDAQ), 28 percent on the New York Stock Exchange (NYSE), and the remaining 13 percent on the American Stock Exchange (ASE). These proportions are similar for Section 404 firms, among which 60 percent, 32 percent, and 8 percent trade on NASDAQ, NYSE, and ASE, respectively. The average market capitalization for Section 302 disclosers ranges from $99 million to $6,286 million and is largest for the NYSE, consistent with NYSE listing requirements. The average market capitalization of firms disclosing weaknesses under Section 404 ranges from $239 million to $2,826 million. Within NASDAQ and ASE, the median market capitalization of 404 disclosers is larger than that of 302 disclosers.

Panel C of Table 1 shows the number of weakness disclosers by industry. Industry groups with the largest representation in the Section 302 sample include Durable Manufacturers (22 percent), Computers (18 percent), Other (14 percent), Services (12 percent), and Retail (11 percent). These are also the most representative groups in the Section 404 sample. Supplemental analyses, not tabulated, show that durable manufacturers, service firms, and firms in the computer industry are disproportionately represented in the disclosure samples relative to the Compustat population.[20]

Because other information may accompany weakness disclosures, we perform a search on Dow Jones News Service to identify other news published in the three-day event window for both the material weakness disclosure (MWD) samples and the control samples. Table 2 presents the results of the contamination analysis for Section 302 MWD firms. Highlighting the need to consider contemporaneous information, we find that nearly half of Section 302 disclosures accompany other news (150 out of 330 Section 302 disclosures). The news items for Section 302 firms are both positive and negative. Earnings news (with a valence

---

[19] To ensure that our sample is consistent with contemporaneous papers (Ashbaugh-Skaife, Collins, and Kinney 2007; Ashbaugh-Skaife, Collins, Kinney, and LaFond 2007; Ashbaugh-Skaife et al. 2008; De Franco et al. 2005; Doyle et al. 2007a, 2007b; Ge and McVay 2005; Hammersley et al. 2008), we match our results to the sample published by *Compliance Week*. We find that our sample has 83 percent overlap with that of *Compliance Week*. As noted in these studies, the event dates supplied by *Compliance Week* often fail to capture the initial date of disclosure. Our sample selection criteria and method of data collection ensure that we accurately capture the date of initial disclosure.

[20] In results not tabulated, we compare Section 302 and Section 404 disclosers (MWD firms) to the Compustat population in the year prior to the fiscal year preceding the disclosure of a material weakness and control firms matched on industry and total assets. We find that MWD firms are larger (in terms of assets, sales and market value of equity), less profitable (lower ROA and higher frequency of losses), and they experience slower sales growth than the median Compustat firm in fiscal 2003; MWD firms have higher accruals (less negative) and are more likely to have reported income-decreasing special items.

TABLE 1
Full Sample of Material Weakness Disclosures

Panel A: Firms Disclosing Internal Control Weaknesses under Sections 302 and 404

Sample	Year	Month	n	Percent
302	2004	January	14	4.24
		February	8	2.42
		March	17	5.15
		April	13	3.94
		May	22	6.67
		June	24	7.27
		July	17	5.15
		August	39	11.82
		September	23	6.97
		October	31	9.39
		November	87	26.37
		December	35	10.61
		Total	330	100.00
404	2005	January	1	0.26
		February	2	0.52
		March	207	54.05
		April	67	17.49
		May	43	11.23
		June	21	5.48
		July	6	1.57
		August	4	1.04
		September	10	2.61
		October	2	0.52
		November	7	1.83
		December	13	3.40
		Total	383	100.00

Panel B: Exchange Listing and Market Capitalization at Disclosure Date

		Market Capitalization (in thousands of dollars)				
Exchange	**n**	**Mean**	**Median**	**Minimum**	**Maximum**	**Std. Dev.**
302 Firms						
NASDAQ	196	451,674	169,294	1,984	9,236,204	987,631
NYSE	92	6,268,350	980,255	26,454	243,213,117	26,804,349
ASE	42	99,401	38,034	4,028	899,368	162,603
Total	330	2,028,458	217,274	1,984	243,213,117	14,362,840
404 Firms						
NASDAQ	229	595,855	279,644	3,677	10,427,193	1,146,622
NYSE	123	2,826,047	909,624	20,134	47,243,396	5,906,186
ASE	31	239,196	175,711	9,168	1,093,439	249,148
Total	383	1,283,210	355,863[a]	3,677	47,243,396	3,615,020

(*continued on next page*)

TABLE 1 (continued)

Panel C: Industry Composition of Section 302 and Section 404 Disclosure Firms

Industry	n	Percent of Firms in the 302 Sample	n	Percent of Firms in the 404 Sample	Percent of Firms in the Compustat Population
Mining and construction 1000–1999, except 1300–1399	7	2.12	11	2.87	2.39
Food 2000–2111	4	1.21	2	0.52	1.80
Textiles, printing, and publishing 2200–2799	11	3.33	10	2.61	2.97
Chemicals 2800–2824, and 2840–2899	8	2.42	6	1.57	1.48
Pharmaceuticals 2830–2836	12	3.64	14	3.66	5.06
Extractive industries 2900–2999, and 1300–1399	10	3.03	10	2.61	2.97
Durable manufacturers 3000–3999, except 3570–3579, 3670–3679	74	22.42	62	16.19	15.69
Computers 7370–7379, 3570–3579, and 3670–3679	59	17.88	79	20.63	12.06
Transportation 4000–4899	26	7.88	27	7.05	5.54
Retail 5000–5999	35	10.61	36	9.40	7.72
Services 7000–8999, except 7370–7379	38	11.52	49	12.79	8.79
Other	46	13.94	77	20.10	33.53
	330	100.00	383	100.00	100.00

[a] The median market capitalization for the 404 firms significantly differs from the median market capitalization for the 302 firms at the 1 percent level for a two-tailed test. The mean market capitalization does not differ significantly between groups.

that depends on expectations) contaminates 85 weakness disclosures. Many other news items relate to mergers, product introductions, FDA approvals, customer service, and support developments and dividend announcements. As discussed in the Section III, we make no predictions as to the valence of this news. However, we do expect negative effects of 26 news items referencing auditor changes (Beneish et al. 2005; Whisenant 2003).[21] Untabulated statistics show that Section 404 firms have proportionately fewer news items than Section 302 firms (37 percent of Section 404 observations are contaminated, relative to 45 percent of Section 302 observations).

Because other news is pervasive in the MWD sample, we compute abnormal returns and changes in costs of capital for the full samples and separately for weakness disclosures that are not released contemporaneously with other information (the "uncontaminated"

---

[21] Additional analysis indicates that 9 of the 26 auditor changes involve an auditor departure while the remaining 17 indicate that a new auditor has been named subsequent to departure. We use this finer partition in the cross-sectional regression.

TABLE 2
Confounding News Events for Firms Disclosing Weaknesses

	Number of Firms	
MWD sample with available CRSP data		330
Less: firms with confounding news items		
Earnings release	85	
Change in auditor	26	
Other news	11	
Change in officer/director	11	
Restatement	8	
Shareholder lawsuit/SEC investigation	5	
Delay in filing	4	
Number of firms with confounding news items		150
Uncontaminated MWD sample with CRSP data		180

To identify confounding news items for the 302 sample, we perform a search on Dow Jones news service to identify other news published in the three-day event window for the disclosure group and three control samples. This process identified 150 disclosing firms with 222 total confounding news items in the event window and 180 uncontaminated observations. The same process identified an uncontaminated industry- and size-matched control sample of 261 firms, an uncontaminated industry- and performance-matched control sample of 249 firms, and an uncontaminated industry- and earnings-quality-matched control sample of 268 firms. The total news items identified (222) differs from the total firms with confounding news items (150) because some firms disclosed multiple news items in the event window. The majority of "other news" relates to new product introductions, mergers and acquisitions, dividend announcements, and financing information.

To identify confounding news items for the 404 sample, we follow the same procedure for the sample of firms filing qualified 404 opinions (n = 383) and for the sample of 302 firms filing unqualified 404 opinions (n = 84). We do not tabulate these results according to the categories above. This process identified an uncontaminated "Good, Bad" sample of 178 firms, an uncontaminated sample of "Bad, Bad" firms of 71 firms, and an uncontaminated "Bad, Good" sample of 46 firms.

samples). We follow the same news search process to identify an uncontaminated industry- and size-matched control sample of 261 firms, an uncontaminated industry- and performance-matched control sample of 249 firms, and an uncontaminated industry- and earnings-quality-matched control sample of 268 firms.

V. EMPIRICAL RESULTS

Control Sample Comparisons

Table 3 reports comparisons of the Section 302 MWD firms to control samples matched on industry and total assets (size-matched), industry and return on assets (performance-matched), and industry and *ELOAD* (earnings-quality-matched). Relying on median comparisons, we find that disclosing firms are generally similar to control firms. Relative to size-matched firms, Section 302 MWD firms have poorer financial performance (lower ROA) and poorer earnings quality (higher *ELOAD*). Relative to performance-matched firms, Section 302 MWD firms have higher leverage. Section 302 MWD firms are similar to earnings-quality-matched firms in terms of size, profitability, leverage, growth, and accruals. This suggests that earnings-quality-matched firms provide a reasonable control group for our analyses.

Consistent with Ashbaugh-Skaife, Collins, Kinney, and LaFond (2007), we find that pre-disclosure, Section 302 MWD firms are riskier as reflected in higher average beta, higher average return variability, and higher dispersion of residual returns. For example, the median beta for Section 302 MWD firms of 1.30 is significantly larger than the corresponding measure in any the control samples, which ranges from 1.00 to 1.08. We also

TABLE 3
Characteristics of Firms Disclosing Material Weaknesses and Firms in Control Samples

Panel A: Comparison of the Section 302 Disclosure Firms to Three Control Samples

	Section 302 Firms			Industry- and Size-Matched Firms			Industry- and Performance-Matched Firms			Industry- and Earnings-Quality-Matched Firms		
	Mean	Median	Std. Dev.	Mean	Median	Std. Dev.	Mean	Median	Std. Dev.	Mean	Median	Std. Dev.
Assets	11,189	310	98,007	4,521	310	41,062	4,842	309	27,999	7,690	348	50,086
Sales	2,496	229	11,811	1,540	235	5,721	3,229	207	18,554	2,542	71	7,650
ROA	(0.05)	0.01	0.22	(0.00)**	0.03**	0.18	(0.02)	0.02	0.17	(0.02)	0.03	0.18
Leverage	.45	.43	.23	0.43	0.41	0.24	0.42**	0.42	0.22	0.43	0.43	0.23
Sales growth	0.14	.06	.48	0.12	0.06	0.37	0.09	0.06	0.36	0.19	0.09	0.56
Total accruals	(0.13)	(0.06)	0.22	(0.13)	(0.05)	0.22	(0.13)	(0.06)	0.22	(0.12)	(0.06)	0.22
Current accruals	(0.05)	(0.00)	0.17	(0.05)	(0.00)	0.17	(0.06)	(0.00)	0.17	(0.05)	(0.00)	0.17
MVE	2,414	236	15,804	2,793	296	12,924	2,617	235	13,907	4,346	379	15,146
ELOAD	0.223	0.088	1.357	0.084*	0.017**	0.437	0.142	0.047	0.606	0.154	0.079	0.462
Total fee (2003)	2,156	660	5,311	1,604	568	4,014	2,155	554	8,409	1,944	641	5,085
Total fee (2004)	3,262	1,097	8,518	2,256	939	5,814	2,302	920**	6,015	2,440	1,028	4,916
Audit fee (2003)	1,376	442	3,551	1,051	398	3,067	1,393	356	6,081	968	411	1,901
Audit fee (2004)	2,616	923	7,016	1,794	717**	4,994	1,697	697**	4,153	1,767	835*	3,194
Med. adj. %Δ total fee	0.414	0.183	0.833	0.208**	0.084**	0.677	0.249**	0.052**	0.728	0.272**	0.067	0.719
Med. adj.%Δ audit fee	0.610	0.373	1.040	0.364**	0.106**	0.854	0.336**	0.006**	0.902	0.427**	0.338	0.859
Beta	1.345	1.304	1.016	1.100**	1.060*	0.887	1.172**	1.076*	0.899	1.121**	1.003**	0.902
Std. of return	0.040	0.036	0.023	0.033**	0.030**	0.018	0.035**	0.031**	0.020	0.034**	0.028**	0.021
Std. of residual	0.039	0.033	0.023	0.031**	0.027**	0.018	0.033**	0.028**	0.020	0.033**	0.026**	0.022

(continued on next page)

TABLE 3 (continued)

Panel B: Comparison of Section 302 Disclosure Firms to Section 404 Disclosure Firms

	MWD 302 Firms (n = 330)			404 Firms (n = 383)			MWD 302 Firms with 404 Information (n = 205)			MWD 302 Firms without 404 Information (n = 125)		
	Mean	Median	Std. Dev.	Mean	Median	Std. Dev.	Mean	Median	Std. Dev.	Mean	Median	Std. Dev.
Assets	11,189	310	98,007	2,782	350	11,371	14,338	604	116,660	5,867	51$^{\#\#}$	53,413
Sales	2,496	229	11,811	898**	237	2,610	2,537	347	8,816	2,427	53$^{\#\#}$	15,669
ROA	(0.05)	0.01	0.22	(0.02)**	0.01	0.18	(0.02)	0.01	0.19	(0.11)$^{\#\#}$	(0.03)$^{\#\#}$	0.26
Leverage	0.45	0.43	0.23	0.40**	0.34**	0.28	0.44	0.42	0.25	.46	0.47	0.20
Sales growth	0.14	0.06	0.48	0.16	0.08	0.38	0.17	0.08	0.44	0.09	0.02$^{\#\#}$	0.54
Total accruals	(0.13)	(0.06)	0.22	(0.15)	(0.05)	0.25	(0.13)	(0.05)	0.23	(0.12)	(0.06)$^{\#}$	0.21
Current accruals	(0.05)	(0.00)	0.17	(0.07)	0.00	0.19	(0.05)	(0.00)	0.17	(0.04)	(0.00)	0.16
MVE	2,414	236	15,804	1,041	374**	2,531	3,517	440	19,719	573$^{\#\#}$	39$^{\#\#}$	3,667
ELOAD	0.223	0.088	1.357	0.044**	(0.011)**	0.699	0.212	0.017	1.665	0.240	0.133$^{\#\#}$	0.560
Total fee (2003)	2,156	660	5,311				2,863	930	6,174	473$^{\#\#}$	227$^{\#\#}$	882
Total fee (2004)	3,262	1,097	8,518				4,355	1,690	9,918	662$^{\#\#}$	293$^{\#\#}$	1,392
Audit fee (2003)	1,376	442	3,551				1,810	581	4,143	342$^{\#\#}$	186$^{\#\#}$	579
Audit fee (2004)	2,616	923	7,016				3,487	1,365	8,182	542$^{\#\#}$	208$^{\#\#}$	1,105
Med. adj. %Δ total fee	0.414	0.183	0.833				0.524	0.348	0.863	0.152$^{\#\#}$	(0.060)$^{\#\#}$	0.696
Med. adj. %Δ audit fee	0.610	0.373	1.040				0.824	0.595	1.064	0.101$^{\#\#}$	(0.206)$^{\#\#}$	0.778
Beta	1.345	1.304	1.016	1.465	1.352	0.878	1.545	1.504	0.986	1.013$^{\#\#}$	0.762$^{\#\#}$	0.981
Std. of return	0.040	0.036	0.023	0.034**	0.032*	0.017	0.034	0.030	0.020	0.051$^{\#\#}$	0.048$^{\#\#}$	0.024
Std. of residual	0.039	0.033	0.023	0.031**	0.029**	0.017	0.032	0.028	0.020	0.050$^{\#\#}$	0.046$^{\#\#}$	0.024

(continued on next page)

TABLE 3 (continued)

**, * Denote instances where the characteristic of the control sample differs significantly from that of the MWD sample at the 5 percent (10 percent) level for a two-tailed test.

##, # Denote instances where the characteristic of the MWD firms without 404 information differs significantly from that of the MWD sample with 404 information at the 5 percent (10 percent) level for a two-tailed test.

Variable Definitions (all data are from 2003 unless otherwise noted):

Assets = Compustat Data Item #6;
Sales = Compustat Data Item #12;
Return on assets (ROA) = Income before extraordinary items (Compustat Data Item #18) divided by assets;
Leverage = current liabilities (Compustat Data Item #5) plus long-term debt (Compustat data item #9) divided by assets;
Sales growth = percentage change in sales;
Total accruals = $-(\Delta AR$[#302] + ΔINV[#303]+ ΔAP[#304] + ΔTAX [#305] + ΔOTH [#307] + *DEP* [#125])/Total assets_{-1}[#6];
Current accruals = Total accruals + DEP[#125]/Total assets_{-1}[#6];
MVE = the product of shares outstanding and price from CRSP at the end of 2003;
Total fee = total fees paid to auditor;
Audit fee = total audit fee paid to auditor;
Beta = market model beta for the year;
Std. of return = standard deviation of returns for the year; and
Std. of residual = standard deviation of residual from market model for the year.

We estimate *ELOAD* over calendar year 2003 for the CRSP population running the following regression:

$$(R_{i,t} - Rf_t) = \alpha_i + \beta_i(R_{M,t} - Rf_t) + s_i SMB_t + h_i HML_t + e_i EQF_t + \varepsilon_{i,t},$$

where:

$R_{i,t}$ = the return to decile *i* for month *t*;
Rf_t = the return on the 30-day T-bill for month *t*;
$R_{M,t}$ = the return on the value-weighted CRSP index for month *t*;
SMB_t = the Fama-French size factor mimicking portfolio;
HML_t = the Fama-French book-to-market factor mimicking portfolio; and
EQF_t = the Ecker et al. (2006) earnings quality factor mimicking portfolio.

find that Section 302 MWD firms experience significantly greater increases in audit and professional fees than firms in two of the three control samples. That is, the median abnormal increase in total (audit) fees for Section 302 MWD firms of 18.3 percent (37.3 percent) is significantly greater than the increase in audit fees and total fees for size-matched firms (8.4 percent and 10.6 percent, respectively) and for performance-matched firms (5.2 percent and 0.6 percent, respectively). These results provide some support for H2 and the notion that MWD firms bear *ex post* greater costs in relation to material internal control weaknesses. However, the impact of these costs on cash flows is relatively small, representing for Section 302 MWD firms, an average 0.6 percent (median 0.4 percent) of pre-disclosure market value of equity.

Panel B of Table 3 presents statistics for the Section 404 MWD sample as well as descriptive statistics for the Section 302 MWD sample conditional on Section 404 filing status (accelerated filers versus non-accelerated filers). Statistics for the Section 302 sample are repeated from Panel A for ease of reference. The second columnar section reveals that Section 404 weakness disclosers are larger than Section 302 disclosers (median market capitalization of \$374 million compared to \$236 million) and have significantly better observable pre-disclosure earnings quality (median *ELOAD* of −0.011 compared to 0.088). In addition, Section 404 disclosers are less risky than Section 302 disclosers in terms of total return variance and dispersion of residual returns (0.032 and 0.029, respectively, compared to 0.036 and 0.033). These findings are consistent with accelerated filers that have not previously disclosed weaknesses having richer information environments.

The third and fourth columnar sections present statistics for Section 302 disclosers, conditional on their Section 404 filing status. The fourth columnar section shows that non-accelerated filers reporting weaknesses under Section 302 are smaller, less profitable, and have poorer earnings quality than accelerated filers reporting weaknesses under Section 302. For example, median market capitalization is \$39 million for non-accelerated filers compared to \$440 million for accelerated filers. Non-accelerated filers are less profitable (median ROA is −0.03 compared to 0.01), and have poorer earnings quality (median *ELOAD* is 0.133 compared to 0.017). Total return variance is also higher (0.048 compared to 0.030), as is residual return variance is higher (0.046 compared to 0.028). Interestingly, the median change in audit fees for non-accelerated filers is negative. Untabulated analyses show that this is due primarily to non-accelerated filers migrating from Big 4 to smaller auditors with lower fee structures. In summary, non-accelerated filers disclosing material weaknesses under Section 302 appear to have less rich information environments based on observable *ex ante* firm and reporting characteristics.

Abnormal Returns to Section 302 Disclosures of Material Weaknesses

Panel A of Table 4 reports statistics describing the market impact of the disclosure of material weaknesses on the full and uncontaminated samples. The cumulative abnormal return is −1.8 percent (p-value < 0.01) for the full sample of 330 observations and the percentage of negative abnormal returns (59.7 percent) is higher than what would be expected by chance. Of the 330 event dates in the sample, 141 are unique (different calendar days) and we address the effect of event day clustering as follows. First, we create an equally weighted portfolio for any calendar day that contains two or more announcing firms, such that the average on a given day becomes one observation in the daily analysis. Second, we compute the mean three-day size-adjusted return over 141 unique days in the sample and find it is reliably negative (−2.31 percent, p-value 0.004), consistent with results using

TABLE 4
Cumulative Abnormal Returns

Panel A: Section 302 Returns for Varying Windows Surrounding the Disclosure Date[a]

	Mean	p-value	Std. Dev.	% Positive
Full sample (n = 330)				
(−1,+1)	−0.018***	0.001	0.096	0.403
Uncontaminated sample (n = 180)				
(−180,−2)	0.025	0.595	0.624	0.460
(−90,−2)	−0.018	0.431	0.308	0.438
(−45,−2)	−0.014	0.371	0.209	0.449
(−30,−2)	−0.011	0.416	0.186	0.420
(−1,+1)	−0.023***	0.001	0.065	0.361
(+2,+30)	0.004	0.766	0.198	0.426
(+2,+45)	−0.004	0.779	0.206	0.466
(+2,+90)	−0.021	0.154	0.192	0.403
(+2,+180)	−0.037	0.296	0.471	0.398

Panel B: Section 302 Three-Day Abnormal Returns Based on Three Control Samples[b]

	n	Mean	p-value	Percent Positive
Return adjusted for industry- and size-matched control firms	330	−0.024***	<.0001	0.376
Contamination:				
Uncontaminated	180	−0.027***	<.0001	0.322
Contaminated	150	−0.022*	0.067	0.440
Return adjusted for industry- and performance-matched control firms	330	−0.016***	0.008	0.439
Contamination:				
Uncontaminated	180	−0.022***	0.000	0.389
Contaminated	150	−0.009	0.415	0.500
Return adjusted for industry- and earnings-quality-matched control firms	330	−0.020***	0.001	0.415
Contamination:				
Uncontaminated	180	−0.025***	0.000	0.383
Contaminated	150	−0.014	0.164	0.453

Panel C: Section 302 Three-Day Abnormal Returns Partitioned by Source, Auditor Quality, and Price[c]

	n	Mean	p-value	Std. Dev.	Percent Positive
Source:					
Current Report	21	−0.039	0.103	0.104	0.333
Annual Report	48	−0.019***	0.010	0.050	0.417
Quarterly Report	87	−0.021***	0.004	0.066	0.356
1933 Act Report	24	−0.020**	0.015	0.037	0.292
Audit Quality:					
Big 4 auditor	130	−0.016***	0.003	0.063	0.415
Small auditor	50	−0.039***	0.000	0.067	0.220
Price:					
Large-price firms	152	−0.019***	0.000	0.058	0.382
Small-price firms	28	−0.041**	0.027	0.092	0.250

(continued on next page)

TABLE 4 (continued)

*, **, *** Denote significance at the 10 percent, 5 percent, and 1 percent level, respectively (two-tailed test).

[a] We obtain stock return data from CRSP and compute size-adjusted cumulative abnormal returns over the windows surrounding the material weakness disclosure by subtracting from individual firm daily returns the concurrent equally weighted market return for other firms in the same market capitalization decile.

[b] We obtain stock return data from CRSP and compute abnormal returns by subtracting the raw return for the control firm (matched on industry and either size, performance or earnings quality) from the raw return for the MWD sample firm.

[c] We obtain stock return data from CRSP and compute size-adjusted cumulative abnormal returns over the three-day window surrounding the material weakness disclosure by subtracting from individual firm daily returns the concurrent equally weighted market return for other firms in the same market capitalization decile.

all observations.[22] Results for the uncontaminated sample are more negative—mean cumulative abnormal returns are −2.3 percent (p-value < 0.01), 63.9 percent of the returns are less than zero, and a clustering analysis for the subsample of uncontaminated observations (on 90 different calendar days) yields similar results (−2.48 percent, p-value 0.001). This suggests that contemporaneous news announcements detailed in Table 3 bias against results in the full sample.

We evaluate return behavior in the periods surrounding the event window by calculating abnormal returns over periods 30, 45, 90, and 180 days before and after the event window. For accelerated filers, calculation of abnormal returns over the post-disclosure period is constrained by subsequent Section 404 disclosures. Therefore, we truncate the Section 302 measurement period two days prior to the date of each firm's Section 404 disclosure. Results show that cumulative abnormal returns do not differ from zero for the Section 302 disclosure firms in any of the intervals up to 180 days prior to the event window. More important, the cumulative abnormal returns are not distinguishable from zero in any of the intervals up to 180 days subsequent to disclosure. This suggests that there is no significant systematic "rebound" effect for disclosures and provides some evidence that the market response to Section 302 disclosures is not transitory.

Panel B of Table 4 presents event-window abnormal returns obtained by subtracting returns of the three control samples described in Table 3 from returns of Section 302 MWD firms. For the purposes of this analysis, disclosing firms and control firms are matched in pairs based on industry and each of size, performance, and *ELOAD*.[23] Abnormal returns are significantly negative across all benchmarks used for computing abnormal returns. This is true for the full sample as well as the uncontaminated samples. For example, full-sample cumulative abnormal returns are −0.024, −0.016, and −0.020 based on the size-matched, performance-matched and earnings-quality-matched samples, respectively. Similarly, uncontaminated-sample abnormal returns are −0.027, −0.022, and −0.025 based on the size-matched, performance-matched, and earnings-quality-matched samples, respectively. The abnormal returns for the 150 contaminated disclosures are typically smaller and only differ statistically from zero only when abnormal returns are based on the size-matched sample (−0.022; p-value < 0.10). This suggests that firms tend to release positive news concurrently with material weakness disclosures.

---

[22] The mean number of announcements per calendar day is approximately 2.4 and the median is 1. The positive skewness is in large part due to three calendar days in which there are 35, 23, and 11 firms disclosing material weaknesses. Clustering can result in a violation of independence across observations and lead to overstated t-statistics. To avoid inflating the significance of our abnormal returns estimates, we rely on two methods of computing standard errors: (1) we use cross-sectional standard errors in the event window and (2) we use the time-series standard deviations of the portfolio average return in the period preceding the event, thus effectively collapsing all sample securities into one security.

[23] Results are robust to using tests of pooled differences rather than matched-pairs. Untabulated results show that abnormal returns for each of the control groups do not differ statistically from zero, except for the size-matched sample, for which the average return is positive.

In Panel C of Table 4 we further partition the uncontaminated sample according to (1) the source of the material weakness disclosure, (2) the quality of the auditor (where auditor size proxies for quality), and (3) whether the price of a firm's stock is below or above three dollars. Results are presented in Panel C of Table 4. The first two partitions follow from our previous discussion. The third is an attempt to assess whether movement between unchanged bid and ask prices for low-priced stocks drives the stock price reaction.[24] Although the abnormal returns range from −1.9 to −3.9 percent, we find no statistically significant differences in market reaction across alternative sources of Section 302 disclosures.

The sample partitioned on auditor quality reveals that firms audited by Big 4 auditors experience significantly less negative stock price reactions (−1.6 percent) than those audited by non-Big 4 auditors (−3.9 percent). This result is consistent with our expectation that high audit quality mitigates negative effects of increased information uncertainty. Our final partition in Panel C indicates that firms experience significantly negative stock price reactions independently of whether we categorize their price as low or high. This suggests that the evidence for the sample as a whole is not driven by either a few low-priced firms or the bid-ask bounce. The average abnormal return of −4.1 percent for the 28 firms with prices below $3 is not statistically different from the −1.9 percent experienced by firms with prices greater than $3, and the return for the latter group is in line with that observed for the sample of uncontaminated disclosures.

Conditional Return Analyses

Our conditional return analyses consider each of three groups of Section 302 disclosers, classified by the quality of their subsequent Section 404 disclosures. We label these groups "bad/bad" "bad/good," and "bad/no-404." The first two groups represent Section 302 disclosers that are accelerated filers for Section 404, and the third group represents non-accelerated filers. The term "bad" refers to disclosure of a material weakness under Section 302 and the term "good" refers to nondisclosure under Section 302. For Section 404 disclosures, the terms "bad" and "good" refer to qualified and unqualified internal control opinions, respectively.

For the uncontaminated sample, Table 5 presents Section 302 abnormal returns conditional on subsequent Section 404 filings and Section 404 abnormal returns conditional on previous Section 302 filings. Returns are also partitioned according to weakness type (i.e., company-level weakness or account-specific weakness). Non-accelerated filers experience significantly negative cumulative abnormal returns (−4.22 percent), and these returns are significantly more negative than the returns experienced by the accelerated filers (p-value < 0.05). Each of the accelerated-filer groups also experiences significantly negative cumulative abnormal returns when the sample is uncontaminated (−1.12 percent for bad/bad firms and −1.07 percent for bad/good firms). Partitioning the Section 302 returns by severity of the weakness shows that for accelerated filers, the return is significantly negative only for firms disclosing more severe company-level weakness. In contrast, the market responds negatively both to company-level and to account-specific weaknesses when disclosed by non-accelerated filers. Overall, these results are consistent with the conjecture

---

[24] Because the choice of $3 as a cut-off is arbitrary, we also experiment with $4 and $5 as cut-offs in our analyses and find similar results.

TABLE 5
Conditional Uncontaminated Cumulative Abnormal Returns for 302 and 404 Event Windows Partitioned on Weakness Severity

	Section 302 Return, Conditional on Subsequent Section 404 Filing Status CAR (−1,+1) Percent Negative (Number of Obs.)			Section 404 Return, Conditional on Previous Section 302 Filing Status CAR (−1,+1) Percent Negative (Number of Obs.)		
	CLW	*ASW*	Both	*CLW*	*ASW*	Both
Panel A: Firms Disclosing Weaknesses under Section 302:						
Bad, Bad	−.0109** 63.0% (n = 31)	−.0114 61.7% (n = 34)	−.0112** 63.0% (n = 65)	.0058 33.3% (n = 21)	−.0051 60.9% (n = 23)	.0001 47.8% (n = 44)
Bad, Good	−.0151** 75.0% (n = 12)	−.0094 63.4% (n = 41)	−.0107** 66.0% (n = 53)	.0186 25.0% (n = 8)	.0019 43.2% (n = 32)	.0049 40.0% (n = 45)
Bad, No 404	−.0328* 64.0% (n = 25)	−.0489*** 66.7% (n = 36)	−.0422*** 65.6% (n = 61)	NA	NA	.0073 58.9% (n = 39)
Panel B: Firms Not Disclosing Weaknesses under Section 302:						
Good, Good	NA	NA	NA	NA	NA	−.0005 47.6% (n = 2013)
Good, Bad	NA	NA	NA	−.0027 54.5% (n = 55)	−.0031 54.0% (n = 123)	−.0030 54.2% (n = 178)

We obtain stock return data from CRSP and compute size-adjusted cumulative abnormal returns over the three-day window surrounding the material weakness disclosure by subtracting from individual firm daily returns the concurrent equally weighted market return for other firms in the same market capitalization decile. Using classification coding supplied by Doyle et al. (2007a), we set CLW equal to 1 if a firm discloses a company-level 302 weakness and equal to 0 otherwise. Similarly, we set ASW equal to 1 if a firm discloses an account-specific weaknesses and equal to 0 otherwise. The "Good, Good" firms represent those firms with no disclosed 302 weakness and an unqualified 404 opinion. The "Good, Bad" firms represent those firms with no disclosed 302 weakness and a qualified 404 opinion. The "Bad, Good" firms represent those firms with a disclosed 302 weakness and an unqualified 404 opinion. The "Bad, Bad" firms represent those firms with a disclosed 302 weakness and a qualified 404 opinion. The "Bad, No 404" firms represent those MWD 302 firms not required to file 404 reports.

that accelerated filers operate in less uncertain information environments, have lower thresholds of disclosure, or both.

Following the same procedure used for the Section 302 disclosures, we compute abnormal returns over the three-day event window surrounding Section 404 announcements. As discussed in Section II, we condition the Section 404 returns on each firm's prior Section 302 disclosure with the expectation that previous weakness disclosure or nondisclosure will influence the market reaction to a subsequent weakness disclosure. Specifically, we expect the most positive (negative) responses to Section 404 disclosures when the prior Section 302 disclosure news is bad (good) because investors are more likely to revise their beliefs when current news contradicts their expectations.

Results are presented in the last three columns of Table 5. At conventional levels, we find no abnormal returns incident to Section 404 disclosures in any partition. Additional analyses, not tabulated, reveal that pooled results are also insignificant, as are results on the full sample. The lack of significance in the good/bad (bad/good) samples are surprising given Ashbaugh-Skaife, Collins, Kinney, and LaFond's (2007) finding that PAR increases (decreases) for these categories of firms.[25]

The lack of market response to Section 404 disclosures can be due to a number of factors. First, as noted above, accelerated filers may have lower thresholds for disclosure because their internal controls are independently audited. In addition, the market's reaction to Section 404 disclosures may be attenuated by firm attributes associated with accelerated filers (e.g., larger size and higher quality auditors). Finally, we may lack power to detect abnormal returns due to noise induced by the clustering of annual report filings coincident with Section 404 disclosures.

Section 404's failure to produce informative disclosures may help to explain why qualified Section 404 disclosures do not explain cost of capital levels (Ogneva et al. 2007). Specifically, Section 404 opinions may not discriminate between firms with serious weaknesses and others. As well, our results show that studies that focus on 404 disclosures exclude the firms for which weakness disclosures are most informative (i.e., non-accelerated Section 302 disclosers).

As a rough measure of the persistence of the negative returns documented in Table 5, we plot in Figure 3 cumulative long-window, equal-weighted portfolio values for the groups represented in Table 5, as well as for a group of earnings-quality matched control firms that do not report material weaknesses under either of Sections 302 or 404 (good/good).[26] The graph reports the value of the portfolios from six months before the Section 302 disclosure to six months after, excluding delisting returns, and assuming that $1 was invested in each portfolio on day −180. Consistent with the results in Table 5, Figure 3 reveals that non-accelerated filers (no-404 firms) experience the most significant, and the only continuing, significant price declines incident to Section 302 announcements.[27] Interestingly, Section 302 disclosures are most informative and are associated with longer-term consequences for firms specifically exempted from Section 404 compliance.

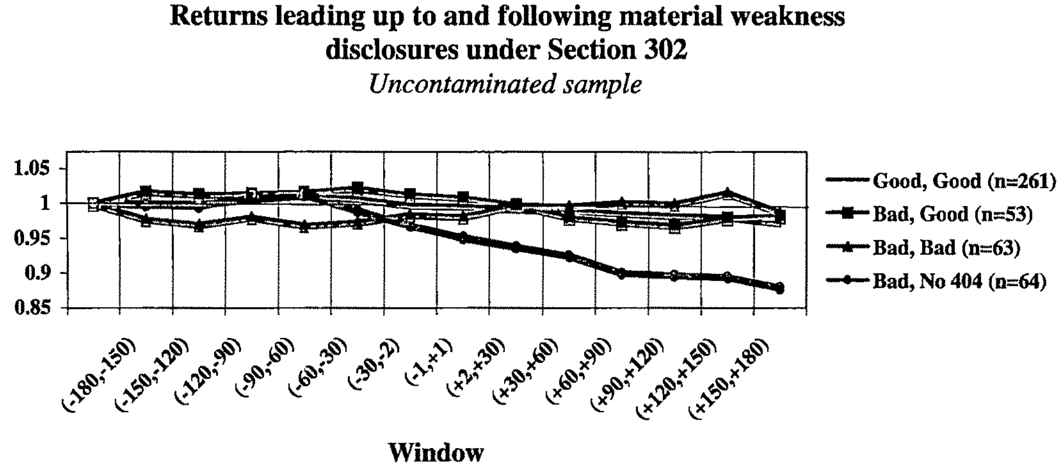
---

[25] If price is unchanged, an increase in PAR implies an increase in expected cash flows that would be inconsistent with identification of an internal control weakness.

[26] The effects of disclosure cannot be well isolated when returns are computed over long windows because of possible information events subject to disclosure. We do not plot the pooled group on the chart. The cumulative abnormal return for the pooled Section 302 disclosure group significantly differs from zero at all intervals in the 180 days subsequent to disclosure.

[27] Specifically, the portfolio value for the "Bad, No 404" group differs from 1 at the 5 percent level (or better for a two-tailed test) at every interval during the 30 days before disclosure through the 180 days after disclosure.

FIGURE 3
Returns Surrounding Section 302 Disclosures



The "Good, Good" firms represent industry- and earnings-quality-matched control firms that do not report qualified 404 opinions. The "Bad, Good" firms represent Section 302 disclosure firms filing subsequent unqualified 404 opinions. The "Bad, Bad" firms represent those Section 302 disclosure firms filing qualified 404 opinions. The "Bad, No 404" firms represent Section 302 disclosure firms not required to file 404 reports.
We plot the cumulative, equal-weighted portfolio values for each group from six months before the 302 disclosure to six months after, assuming an initial investment of $1 on day −180.
The portfolio value for the "Bad, No 404" group differs from 1 at the 5 percent level (or better for a two-tailed test) at every interval during the 30 days before disclosure through the 180 days after disclosure.

Cross-Sectional Regression Results

In Table 6, we present five estimations of the regression model described in Equation (1). Recognizing that ordinary least squares (OLS) may not properly account for the effect of clustering in time and lead to overstated statistical significance in the regression analyses, we report two-tailed statistical significance using both OLS standard errors and robust clustered standard errors (Huber 1967). Procedurally, we implement Huber (1967) using PROC GENMOD in SAS. The first two estimations differ across alternative audit quality variables and are based on 330 observations. The next two estimations include *NO404REQ* and *CLW* (n = 330), while the final two include *ELOAD* (n = 324) and both *ELOAD* and *FEECHGE* (n = 247), causing reductions in sample size.

In all estimations we find that the auditor quality variables have significant, positive coefficients and that the source and nature of the material weakness disclosure does not affect its market impact. The result on auditor quality suggests that the shock to financial reporting credibility is attenuated by the presence of a high-quality auditor, even though the disclosures themselves are unaudited. This is consistent with prior research that suggests that auditor characteristics affect the perceived credibility of financial reporting systems. The coefficient on *RESIGN* is not generally significant across the estimations.

Based on the consistent significance of *CONTAMINATION*, we conclude that nonaudit contaminating information attenuates the adverse market reaction, consistent with intentional or unintentional coupling of the weakness disclosure with positive other information.

TABLE 6
Cross-Sectional Regression of Cumulative Abnormal Returns

$$CAR_{IT} = \alpha_0 + \alpha_1 AQ_{i,t} + \alpha_2 RESIGN_{i,t} + \alpha_3 AUDCHGE_{i,t} + \alpha_4 CONTAMINATION_{i,t} + \alpha_5 CONTEXT_{i,t} + \alpha_6 NO404REQ_{i,t}$$
$$+ \alpha_7 CLW_{i,t} + \alpha_8 ELOAD_{i,t} + \alpha_9 FEECHGE_{i,t} + \varepsilon_{i,t}.$$

		Model 1 (n = 330)			Model 2 (n = 330)			Model 3 (n = 330)			Model 4 (n = 324)			Model 5 (n = 247)		
	Pred.	Coeff.	pr > |t|	pr > |z|	Coeff.	pr > |t|	pr > |z|	Coeff.	pr > |t|	pr > |z|	Coeff.	pr > |t|	pr > |z|	Coeff.	pr > |t|	pr > |z|
Intercept	−	−0.021	0.187	0.161	−0.026	0.118	0.126	−0.029	0.095	0.100	−0.033	0.063	0.069	−0.045	0.023	0.034
AQ																
BIG	+	0.025	0.046	0.078	—	—	—	—	—	—	—	—	—	—	—	—
NONSPEC_BIG	+	—	—	—	0.029	0.040	0.066	0.030	0.035	0.056	0.034	0.020	0.037	0.042	0.011	0.028
INSPEC	+	—	—	—	0.031	0.037	0.052	0.031	0.036	0.050	0.033	0.028	0.042	0.048	0.005	0.013
RESIGN	−	−0.072	0.037	0.196	−0.067	0.055	0.228	−0.067	0.054	0.228	−0.066	0.057	0.233	−0.037	0.333	0.127
AUDCHGE	?	−0.038	0.118	0.245	−0.038	0.118	0.251	−0.038	0.118	0.246	−0.038	0.123	0.253	−0.033	0.228	0.373
CONTAMINATION	?	0.027	0.014	0.026	0.027	0.014	0.027	0.027	0.014	0.026	0.027	0.014	0.025	0.025	0.040	0.050
CONTEXT=CURR	−	−0.011	0.522	0.559	−0.013	0.452	0.485	−0.012	0.470	0.505	−0.011	0.522	0.556	0.000	0.990	0.991
CONTEXT=REG	−	−0.020	0.135	0.150	−0.020	0.133	0.152	−0.020	0.132	0.153	−0.002	0.189	0.215	−0.021	0.150	0.209
CONTEXT=K	−	−0.002	0.902	0.890	−0.002	0.933	0.926	−0.003	0.886	0.874	−0.002	0.913	0.905	0.003	0.906	0.904
NO404REQ	−	−0.026	0.027	0.057	−0.026	0.023	0.050	−0.026	0.023	0.049	−0.026	0.024	0.050	−0.015	0.291	0.387
CLW	+	—	—	—	—	—	—	0.007	0.540	0.556	0.009	0.427	0.441	0.008	0.488	0.499
ELOAD	+	—	—	—	—	—	—	—	—	—	0.000	0.924	0.864	0.024	0.060	0.057
FEECHGE	−	—	—	—	—	—	—	—	—	—	—	—	—	0.004	0.608	0.585
Adj. R²			7.36%			7.40%			7.22%			7.38%			6.66%	

Variable Definitions:

BIG = 1 if a Big 4 accounting firm audits the firm;

NONSPEC_BIG = 1 if a non-industry specialist Big 4 accounting firm audits the firm;

INDSPEC = 1 if an industry specialist audit firm audits the firm, following Craswell et al. (1995). We use a threshold of 20 percent of industry sales to designate industry specialists. If an industry has less than 30 firms, we designate no specialists for that industry;

RESIGN = 1 if the current auditor resigns in the event window;

AUDCHGE = 1 if the firm appoints a new auditor in the event window;

CONTAMINATION = 1 if any of the other confounding news events (described in Table 2) occur during the event window;

CONTEXT = three indicator variables designed to control for the source of the material weakness disclosure.

CLW = 1 if a firm discloses a company-level weakness and equal to 0 otherwise, using coding supplied by Doyle et al. (2007a); and

FEECHGE = the median-adjusted percentage change in total fee from 2003 to 2004.

We define *ELOAD* in Table 3.

Again consistent with accelerated filers operating in richer information environments, the coefficient on *NO404REQ* exhibits a significantly negative coefficient. In one of the two estimations, *ELOAD* exhibits the predicted positive coefficient. The coefficient on the *FEECHGE* variable is not distinguishable from zero at conventional levels after controlling for other factors. This suggests either that this is a poor proxy for expected incremental fees, or that this variable exhibits insufficient cross-sectional variation within the MWD sample.[28]

In summary, we find that material weakness disclosures negatively affect the stock prices of disclosing firms, consistent with H1. At one extreme, these disclosures can be purely qualitative (i.e., have no cash flow effect), suggesting that investors increase the rate at which they discount the firm's future cash flows. At the other extreme, the disclosures can represent purely a cash flow effect, indicating that investors are reacting only to expected remediation and other costs. Because the returns analysis cannot distinguish between these two possibilities, we next investigate changes in the cost of capital subsequent to disclosure.

Changes in Cost of Capital Subsequent to Disclosure

We collect analyst forecast data from I/B/E/S for firms disclosing weaknesses and industry- and earnings-quality-matched controls. There are sufficient data to estimate costs of capital for 75 of the Section 302 observations and 84 Section 302 control firms. There are sufficient data to estimate costs of capital for 67 firms initially disclosing weaknesses under Section 404 (good/bad firms), 21 firms repeating Section 302 weaknesses in their Section 404 reports (bad/bad firms), and 32 firms remediating weaknesses (bad/good firms). Untabulated analyses reveal that firms surviving the I/B/E/S data screens are significantly larger. Only 6 of the 75 Section 302 firms with sufficient data to estimate cost of capital are non-accelerated filers. For Section 302 disclosers, announcement period abnormal returns are significantly negative for firms with and without forecast data, but the group of firms without sufficient data experience negative returns in excess of the group with analyst data. The larger size and less negative market response for the group of firms on which we perform cost of capital computations suggests that data limitations bias *against* finding evidence of increasing costs of capital.

Panel A of Table 7 presents descriptive statistics for our cost of capital estimates for Section 302 disclosers and the earnings-quality-matched control firms. Control firms surviving the data screens have larger total assets (mean of $20,152 million compared to $18,768 million) and larger market capitalization (mean of $6,920 million compared to $4,797 million). Profitability, *ELOAD*, return variability, and residual return variability do not differ significantly between disclosing firms and the control groups, but Section 302 firms have higher betas than control firms.[29]

Consistent with H2, results show that Section 302 disclosers experience a less positive forecast revision than control firms. Specifically, the mean (median) total earnings per share revision for the control sample is $0.267 ($0.248) while the mean total earnings per

---

[28] We obtain similar results when *FEECHGE* is measured as audit fees, total fees, and when the median adjustment is drawn from the population with available data or from each of the control samples. Also, we obtain similar results when controlling for industry and month effects in our estimations, though both *RESIGN* and *ELOAD* become insignificant at conventional levels.

[29] In Table 3, Panel A, we reported that the full sample of 330 Section 302 firms had substantially higher market-based risk measures than control firms. This is not the case for the subsample of MWD and control firms for which data are available to estimate the cost of capital. We attribute this apparent inconsistency to the fact that the cost if capital subsample consists of larger firms with greater analyst coverage.

TABLE 7
Cost of Capital Surrounding 302 Disclosures

	MWD Firms (n = 75)					Industry- and Earnings-Quality-Matched Firms (n = 84)					Differences	
	Mean	Pr > |t|	Med.	Min	Max	Mean	Pr > |t|	Med.	Min	Max	Mean	Med.
Panel A: Descriptive Statistics[a]												
ELOAD	0.006		−0.029	−0.585	0.811	0.068		0.010	−0.640	1.309		
Assets	18,768		654	43	1,264,032	20,152		1,729	19	736,445		***
MVE	4,797		707	114	43,369	6,920		1,426	15	70,428		**
ROA	0.026		0.026	−0.374	0.244	0.036		0.038	−0.501	0.209		
Beta	1.538		1.505	0.315	3.268	1.357		1.275	0.228	2.786	**	***
Std. of return	0.028		0.026	0.012	0.061	0.026		0.022	0.011	0.054		
Std. of residual	0.025		0.023	0.010	0.061	0.023		0.021	0.009	0.054		
Price change	−0.009	0.220	−0.010	−0.207	0.152	0.015	0.013	0.012	−0.156	0.165	***	**
EPS revision	0.040	0.600	0.016	−2.181	2.380	0.267	0.002	0.248	−3.387	2.881	***	***
% < 0 EPS revision	44.0%					21.4%						***
ltgb	0.178		0.137	0.026	1.213	0.148		0.133	0.010	0.545		
ltga	0.178		0.145	0.026	1.257	0.151		0.133	0.010	0.545		
ltg_revision	0.001	0.894	0	−0.108	0.180	0.002	0.168	0	−0.050	0.110		
% < 0 ltg_revision	62.5%					48.7%						***

(*continued on next page*)

TABLE 7 (continued)

Panel B: Full Samples, Estimates Based on the Average Price of Firm in the Five Days Leading Up to (Following) the Three-Day Event Window[a]

r_ctc	0.0048	0.055	0.0036	−0.0456	0.0991	0.0055	0.001	0.0039	−0.0367	0.0538		
r_eastonc	0.0170	0.001	0.0129	−0.0995	0.1790	−0.0007	0.874	−0.0025	−0.1206	0.1740	***	***
r_gmc	0.0093	0.003	0.0064	−0.0750	0.0981	0.0003	0.924	−0.0022	−0.0716	0.1117	**	***
r_glsc	0.0053	0.005	0.0046	−0.0424	0.0618	0.0041	0.003	0.0038	−0.0409	0.0455		
mean_rb	0.0891		0.0886	0.0453	0.1339	0.0851		0.0845	0.0389	0.1436		
mean_ra	0.0982		0.0959	0.0378	0.1897	0.0874		0.0869	0.0438	0.1770	***	***
mean_rc	0.0091	0.000	0.0077	−0.0478	0.0809	0.0023	0.240	0.0010	−0.0414	0.0673	**	**

, * Denote instances where the characteristic of the control sample differs significantly from that of the MWD sample at the 10 percent, 5 percent, and 1 percent level, respectively (two-tailed test).

[a] Variable Definitions:

We define *ELOAD, Assets, MVE,* and *Return on Assets* in Table 3. *EPS revision* combines the change in analysts' fiscal year 1 and fiscal year 2 earnings per share (EPS) estimates over the event window. All EPS revisions are nonzero. The *percentage of negative EPS revisions* is the number of EPS revisions less than zero divided by the sample size. The predicted long-term growth rate forecast is measured before the event window (*ltgb*) and after the event window (*ltga*). The revision in long-term growth forecasts (*ltg_revision*) is the average difference between *ltga* and *ltgb*, including firms for which the revision is zero. Approximately 42 percent of firms have zero long-term growth revisions. The percentage of negative long-term growth revisions is the number of nonzero long-term growth forecast revisions that are negative divided by nonzero revisions. We derive estimates of the *change* in cost of capital (*r_ctc, r_eastonc, r_gmc,* and *r_glsc*) following the models described in Dhaliwal et al. (2006, Appendix). Specifically, *r_ctc* follows Claus and Thomas (2001), *r_eastonc* follows Easton (2004), *r_gmc* follows Gode and Mohanram (2003), and r_glsc follows Gebhardt et al. (2001). We calculate mean cost of capital estimates before (*mean_rb*) and after (*mean_ra*) disclosure as well as the mean change in cost of capital (*mean_rc*) for firms with cost of capital estimates available for all four of the models.

[b] Variable Definitions:

We derive estimates of the *change* in cost of capital (*r_ctc, r_eastonc, r_gmc,* and *r_glsc*) following the models described the Appendix. Specifically, *r_ctc* follows Claus and Thomas (2001), *r_eastonc* follows Easton (2004), *r_gmc* follows Gode and Mohanram (2003), and *r_glsc* follows Gebhardt et al. (2001). We calculate mean cost of capital estimates before (*mean_rb*) and after (*mean_ra*) disclosure as well as the mean change in cost of capital (*mean_rc*) for firms with cost of capital estimates available for all four of the models.

share revision for disclosing firms does not differ statistically from zero. As well, nearly twice as many disclosing firms than control firms experience negative EPS revisions (44 percent of disclosing compared to 21 percent of control firms). Disclosers concurrently experience an abnormal mean price change of −$0.024 per share (−$0.009 for disclosers compared to $0.015 for the control group).

In the context of an abnormal price decline and an abnormal decrease in expected future cash flows, the change in disclosing firms' implied costs of capital is indeterminate. Given an abnormal decrease in cash flows, a sufficiently small (large) price decline will result in a decrease (increase) in implied equity cost of capital. To determine whether the required rate of return changes, we solve for the change in equity cost of capital implied by observed price and cash flow changes. The alternative to finding an abnormal change in the cost of capital for disclosing firms is that abnormal price declines documented in the previous section reflect only negative abnormal forecast revisions.

Panel B of Table 7 presents levels and changes in implied costs of equity capital. As suggested by Dhaliwal et al. (2006) and as described more fully in the Appendix, we implement four alternative discounted cash flow models and use the mean in our tests. Disclosing firms experience a mean abnormal increase in cost of capital of 68 basis points, consistent with an adverse impact on firms' financial reporting credibility (91 basis points for disclosing firms, compared to 23 basis points for the control firms). These results suggest that abnormal price increases experienced by disclosing firms are attributable to negative abnormal forecast revisions and positive abnormal increases in required rates of return. The modest, but statistically significant changes are consistent with H3. The results presented in Panel B of Table 7 use price changes over the five days before and after the event window. We replicate our calculations for longer price-change windows (up to 180 days before and after the event window) and inferences are unchanged.

We follow the same procedure to calculate cost of capital changes surrounding firms' disclosures of weaknesses under Section 404. We do not tabulate these analyses because results are comprehensively insignificant at conventional levels. Specifically, we find no evidence of abnormal forecast revisions or cost of capital changes for any partition (good/bad, bad/good, bad/bad). This is consistent with our abnormal returns analysis for Section 404.

We evaluate the robustness of our findings by performing a series of supplemental analyses. We do not tabulate these results. First, we replicate our analyses on control groups comprised of performance- and size-matched controls. Inferences are unchanged. Second, for the five-day price change windows, we replicate our analysis on uncontaminated observations with sufficient data for calculating cost of capital. Inferences are unchanged. Third, because Easton and Monahan (2005) identify *ex ante* criteria that are associated with higher reliability of the implied cost of capital metric, we partition our sample according to their criteria and results remain robust.[30] Fourth, Hail and Leuz (2006) suggest weighting cost of capital estimates based on *ex post* forecast accuracy deflated by price. We weight each observation by the inverse of this measure. If the absence of timely analyst revisions leads to greater measurement error, then this analysis gives more weight to more accurate forecasts, and reduces the influence of estimates with greater measurement error. However, this procedure also potentially underweights observations for which information uncertainty is highest. Consistent with this conjecture, we find that the accuracy-weighted results reveal a smaller, but still significant, increase in cost of capital. Fifth, we collect data from Value Line to replicate the measure used in Ashbaugh-Skaife, Collins, Kinney, and LaFond (2007). We find that results for firms disclosing weaknesses depend on the control group used to calculate abnormal returns. Specifically, disclosing firms experience an abnormal increase in cost of capital when the rest of the Value Line population is used as a control group, but not when industry and earnings-quality-matched firms are used as a control group.

---

[30] Specifically, Easton and Monahan (2005) find that implied cost of capital reliability is highest in the lower tercile of forecasted long-term growth (ltg). Approximately 29 percent of our 75 cost of capital sample firms fall into the lower third of Easton and Monahan's (2005) ltg distribution. We find that the magnitude of the cost of capital change is significant and larger for firms falling into the lower ltg partition (123 basis points for the bottom tercile, compared to 91 basis points for the full disclosure sample). Untabulated statistics comparing the upper two thirds of the sample to the lower one third reveal that the variance in implied cost of capital increases as we move from lower to higher ltg, and that the mean change in implied cost of capital is not significant in the highest tercile for sample firms. This pattern of results is consistent with Easton and Monahan (2005)'s hypothesis that implied cost of capital measures are noisier for firms with higher forecasted ltg rates.

Finally, we evaluate whether MWD announcements are associated with increases in bid-ask spreads. We measure the structural shift in the bid ask spread by adapting the methodology of Affleck-Graves et al. (2000) to our setting.[31] We find a significantly positive abnormal spread for Section 302 weakness disclosers over the event window.

VI. SUMMARY AND CONCLUSION

We evaluate the capital market effects of material weakness disclosures under Sections 302 and 404 of the Sarbanes-Oxley Act (SOX). In doing so, we supply evidence that these disclosures inform investors about the financial reporting quality of disclosing firms. We draw inferences from a sample of 330 firms making Section 302 disclosures and 383 firms disclosing weaknesses under Section 404. We find that firms experience significantly negative abnormal returns when making unaudited disclosures under Section 302. However, we detect no market response to disclosures made under Section 404. These results are consistent with the hypothesis that Section 404 disclosures may reflect a lower materiality threshold for disclosure (Doyle et al. 2007b). The results are also consistent with our conjecture that firms required to make Section 404 disclosures operate in richer information environments that attenuate the market response to weakness disclosure. Indeed, we find that the market response to Section 302 disclosures is significantly smaller for firms required to subsequently file Section 404 reports (accelerated filers). The cumulative abnormal return to disclosure is −4.22 percent for non-accelerated filers and −1.10 percent for accelerated filers.

Our cost-of-capital analyses reveal similar findings. For Section 302 disclosures, we find that firms experience negative abnormal revisions in forecasted earnings per share and positive abnormal increases in implied cost of capital. Yet, we find no evidence that cost of capital increases when firms disclose weaknesses under Section 404. Our finding that the market response to weakness disclosure depends on firms' accelerated filing status may

---

[31] We thank the referee for this suggestion. We measure the abnormal spread for disclosing firms on the event date by estimating the following regression equation with percentage spread (PS) as the dependent variable:

$$\begin{aligned} PS = {} & \alpha + \beta_1 PRICE + \beta_2 VOLUME + \beta_3 EVENT + \beta_4 MWD + \beta_5 (EVENT \times MWD) \\ & + \beta_6 OTHER\ NEWS. \end{aligned}$$

Affleck-Graves et al. (2000) include *PRICE* and *VOLUME* to control for the inventory and order cost effects reflected in the observed spread. The remaining variables are indicator variables: (1) *EVENT* equals 1 if the spread is measured in the event window and 0 otherwise; *MWD* is coded 1 for disclosing firms and 0 for control firms; *OTHER NEWS* equals 1 when there are other news releases during the event window. We are interested in β_5, the coefficient on the interaction between *MWD* and *EVENT*. Specifically, we estimate that β_5 equals .24, (p-value = 0.006) suggesting that Section 302 disclosing firms experience an abnormally higher spread relative to control firms over the event window.

help to explain why weakness disclosures under Section 404 have little explanatory power for firms' risk levels (Ogneva et al. 2007). First, Section 404 disclosures may not identify firms with weaknesses that market participants consider material. Second, classifying firms' risk as high or low by Section 404 disclosures alone eliminates from consideration non-accelerated filers. These firms likely have the highest pre-disclosure information uncertainty. Indeed, our results have implications for the debate about whether a scaled securities regulation system for smaller public companies should be implemented (SEC Advisory Committee on Smaller Public Companies 2006): had Section 302 exemptions mirrored Section 404 exemptions in terms of firm size, the material weakness disclosures would, at best, be weakly informative.

Our study contributes to three important streams of research: (1) the effect of information uncertainty on the pricing of equity securities, (2) the effect of auditor attestation on the credibility of reporting, and (3) the capital market effects of the Sarbanes-Oxley Act. Our analysis, however, is subject to a number of limitations. First, we assume that analysts' forecasts are timely and reflect the market's expectation of future earnings. Second, our study relies on observed stock prices and analysts' expectations of future earnings to solve for the implied cost of capital. Because of these assumptions, any bias in our proxies may lead to a corresponding bias in our calculated change in cost of capital. As well, our search for contaminating news events may not fully reveal all potentially confounding news affecting values. Finally, our analyses are performed using samples limited by data availability. This may cause our tests to lack sufficient statistical power to detect relations in the data.

APPENDIX
COST OF EQUITY ESTIMATES

We compute implied cost of equity capital using methods derived by Gebhardt et al. (2001), Claus and Thomas (2001), Gode and Mohanram (2003), and Easton (2004). We average the results obtained under each of the methods to obtain the mean change in equity cost of capital as established in Dhaliwal et al. (2006).

The methods of Gebhardt et al. (2001) and Claus and Thomas (2001) are special cases of the residual income model and the methods in Gode and Mohanram (2003) and Easton (2004) are special cases of the abnormal growth valuation model. Each of the methods uses analysts' forecasts of earnings to proxy for expected earnings in the near term. To compute the median forecast prior to the event window for each future forecast period, we collect the oldest forecast by each analyst within the 180-calendar-day window that precedes the first day of the event window. To compute the median forecast subsequent to the event window for each future forecast period, we collect the most recent forecast by each analyst within the 180-calendar-day window that follows the last day of the event window.[32] Our inferences are unchanged when we collect analyst forecasts 90 days before and after the event date, or when we use the most recent forecast in the pre-event period. Consistent with Hribar and Jenkins (2004), we adjust I/B/E/S forecasts that shift by one year after the earnings are released by using the actual earnings per share (EPS) following the earnings release.

In general, we compute the dividend payout from Compustat as dividends per share divided by EPS before extraordinary items. If these data items are missing, then we use common dividends divided by income before extraordinary items. Following prior research, we make the following adjustments. If EPS is negative, then we use a return on assets of 6 percent to proxy for long-term EPS in the denominator. We trim observations with payout ratios greater than or equal to 1. Results do not differ when we set observations for which the dividend payout exceeds 100 percent to the median industry payout ratio. To facilitate comparability to other research, we tabulate results using a constant payout rate of 0.5 in the CT model. Overall, we find that our results are relatively insensitive to alternative assumptions we make about long-term dividend payouts. This is most likely because we base our inferences on changes in equity cost of capital (ECC), rather than levels, and dividends exhibit little time-series variability.

---

[32] When a particular analyst's forecast is the same before and after the event window, we have no way of knowing whether the analyst actively intends the amount to be the same, or has passively failed to revise his or her forecast. Therefore, we do not omit these analysts from the analysis. However, inferences are unchanged when we eliminate the forecasts of analysts who forecast the same EPS before and after the event window.

Estimation procedures for future book values and for earnings beyond the horizon for which forecasts are available are particular to each model. We discuss these under the headings for each model, below.

Gebhardt et al. (2001) (GLS)

GLS estimate cost of capital using the following empirical specification:

$$P_t = B_t + \frac{FROE_{t+1} - r_{GLS}}{1 + r_{GLS}} B_t + \frac{FROE_{t+2} - r_{GLS}}{(1 + r)^2} B_{t+1} + TV \qquad \text{(A1)}$$

where:

B_t = book value per share at time t;
$FROE_{t+i}$ = forecasted return on equity (ROE) for period $t + i$; $FROE_{t+1}$ equals the I/B/E/S one-year-ahead forecasted earnings per share (eps) divided by B_t; $FROE_{t+2}$ equals the I/B/E/S two-year-ahead forecasted eps divided by B_{t+1}; $FROE_{t+3}$ equals eps_{t+3} divided by B_{t+2}, and so on;
B_{t+i} = year $t + i$ book value per share; $B_{t+i} = B_{t+i-1} + eps_{t+i} - dps_{t+i}$; when unavailable, we set eps_{t+3} equal to $eps_{t+3}*(1 + ltg)$, where ltg equals the forecasted long-term growth rate of earnings at $t + 2$;[33] forecasted dividends per share (dps_{t+i}) equals eps_{t+i} times the year t actual dividend payout ratio;
TV = terminal value, where we calculate TV as:

$$TV = \sum_{i=3}^{T-1} \frac{FROE_{t+i} - r_{GLS}}{(1 + r_{GLS})^i} B_{t+i-1} + \frac{FROE_{t+T} - r_{GLS}}{r_{GLS}(1 + r)^{T-1}} B_{t+T-1}.$$

We assume each firm's $FROE$ reverts to its industry median ROE over years $t + 4$ to $t + T$ in a linear manner. We calculate r_{GLS} for $T = 12$. We use the 48 Fama and French (1997) industries and calculate median ROE as the ten-year moving median ROE for all firms in the same industry, excluding firms for which income before extraordinary items is less than zero; and

---

[33] When long-term growth (ltg) is unavailable, Dhaliwal et al. (2006) estimate the ltg rate as $(EPS_{t+2}/EPS_{t+1})-1$ for firms with positive values of EPS_{t+i}. We find that ltg is unavailable for approximately 30 percent of sample observations that otherwise have sufficient data, and we estimate ltg, consistent with Dhaliwal et al. (2006) for these firms. We check robustness to this procedure by performing two alternative sets of analyses. First, we delete observations for which ltg is unavailable, resulting in a smaller sample size. Second, we set missing ltg rates equal to the median ltg rate for other firms of similar industry and size (by size decile). Inferences are unchanged when we use either of these alternative procedures. Results are tabulated using the method of Dhaliwal et al. (2006).

P_t = price at time t.

Claus and Thomas (2001) (CT)

CT estimate cost of capital using the following empirical specification:

$$P_t = B_t + \frac{AE_{t+1}}{(1 + r_{CT})} + \frac{AE_{t+2}}{(1 + r_{CT})^2} + \frac{AE_{t+3}}{(1 + r_{CT})^3} + \frac{AE_{t+4}}{(1 + r_{CT})^4} + \frac{AE_{t+5}}{(1 + r_{CT})^5} + \frac{AE_{t+5}(1 + g_{ae})}{(1 + r_{CT})^6} \tag{A2}$$

where:

AE_t = expected abnormal earnings for year t, calculated as $eps_{t+i} - r_{CT}*(B_{t+i-1})$; when no forecast is provided for $t + 3$ through $t + 5$, we multiply the prior-year forecast by $(1 + ltg)$;[34] $B_{t+i} = B_{t+i-1} + eps_{t+i} - dps_{t+i}$; we set dps_{t+i} equal to the assumed dividend payout ratio (0.5) times eps_{t+i};[35] and

g_{ae} = growth rate of abnormal earnings beyond $t + 5$, which we set equal to the yield on ten-year treasury bonds minus 3 percent.

Gode and Mohanram (2003) (GM)

GM estimate cost of capital using the following empirical specification:

$$r_{GM} = A + \sqrt{A^2 + \frac{eps_1}{P_0} * \left(\frac{eps_2 - eps_1}{eps_1} - (\gamma - 1)\right)} \tag{A3}$$

where:

$$A = \frac{1}{2}\left((\gamma - 1) + \frac{dps^1}{P^0}\right);$$

eps_1 = forecasted earnings per share in year 1;
eps_2 = forecasted earnings per share in year 2;
dps_1 = eps_1 multiplied by the dividend payout ratio at $t = 0$; we set observations for which the dividend payout exceeds 100 percent to the median industry payout;
$(\gamma - 1)$ = growth rate in abnormal earnings beyond the forecast horizon minus 1; we set γ equal to the yield on ten-year treasury bonds minus 3 percent; and
P_t = price at time t.

Easton (2004)

Easton (2004) estimates cost of capital using the following empirical specification:

$$r_{EASTON} = \sqrt{\left(\frac{eps_2 - eps_1}{P_t}\right)} \tag{A4}$$

---

[34] See footnote 33.

[35] Inferences are unchanged when the dividend payout rate in year t is used in lieu of 0.5.

where:

eps_1 = forecasted earnings per share in year 1;
eps_2 = forecasted earnings per share in year 2; and
P_t = price at time t.

REFERENCES

Affleck-Graves, J., C. Callahan, and R. Ramanan. 2000. Detecting abnormal bid-ask spread: A comparison of event study methods. *Review of Quantitative Finance and Accounting* 14: 45–65.

Arbel, A., and P. Strebel. 1982. The neglected and small firm effects. *The Financial Review* 17: 201–218.

Ashbaugh-Skaife, H., D. Collins, and W. Kinney. 2007. The discovery and reporting of internal control deficiencies prior to SOX-mandated audits. *Journal of Accounting and Economics* 44: 166–192.

———, ———, ———, and R. LaFond. 2007. The effect of internal control deficiencies on firm risk and cost of equity capital. Available at: http://ssrn.com/abstract=896760.

———, ———. ———, and ———. 2008. The effect of SOX internal control deficiencies and their remediation on accrual quality. *The Accounting Review* 83 (1): 217–250.

Barry, C. B., and S. J. Brown. 1984. Differential information and the small firm effect. *Journal of Financial Economics* 13: 283–294.

Becker, C., M. L. DeFond, J. Jiambalvo, and K. R. Subramanyam. 1998. The effect of audit quality on earnings management. *Contemporary Accounting Research* 15: 1–24.

Bedard, J. 2006. *Sarbanes-Oxley Internal Control Requirements and Earnings Quality.* Available at: http://ssrn.com/abstract=926271.

Beneish, M. D., and J. C. Gardner. 1995. Information costs and liquidity effects in the Dow Jones Industrial Average listing changes. *Journal of Financial and Quantitative Analysis* 30: 135–157.

———, P. E. Hopkins, I. Ph. Jansen, and R. Martin. 2005. Do auditor resignations reduce uncertainty about the quality of firms' financial reporting? *Journal of Accounting and Public Policy* 24: 357–390.

———, M. B. Billings, and L. D. Hodder. 2008. Internal control weaknesses and information uncertainty. *The Accounting Review* 83 (3): 665–703.

Bockus, K., and F. Gigler. 1998. A theory of auditor resignation. *Journal of Accounting Research* 36: 191–197.

Botosan, C. 1997. Disclosure level and the cost of equity capital. *The Accounting Review* 72: 323–349.

Bowen, R., L. DuCharme, and D. Shores. 1995. Stakeholders' implicit claims and accounting method choice. *Journal of Accounting and Economics* 20: 255–295.

Claus, J., and J. Thomas. 2001. Equity premia as low as three percent? Evidence from analysts' earnings forecasts for domestic and international stock markets. *Journal of Finance* 56: 1629–1666.

Craswell, A., and S. Taylor. 1991. The market structure of auditing in Australia: The role of industry specialization. *Research in Accounting Regulation* 5: 55–77.

———, J. Francis, and S. Taylor. 1995. Auditor brand name reputations and industry specializations. *Journal of Accounting and Economics* 20: 297–322.

DeFond, M., and J. Francis. 2005. Audit research after Sarbanes-Oxley. *Auditing: A Journal of Practice & Theory* 24: 5–30.

De Franco, G., Y. Guan, and H. Lu. 2005. *The Wealth Change and Redistribution Effects of Sarbanes-Oxley Internal Control Disclosures.* Available at: http://ssrn.com/abstract=706701.

Dhaliwal, D., S. Heitzman, and O. Li. 2006. Taxes, leverage, and the cost of equity capital. *Journal of Accounting Research* 44: 1–34.

Doyle, J., W. Ge, and S. McVay. 2007a. Determinants of weaknesses in internal control over financial reporting. *Journal of Accounting and Economics* 44: 193–223.

———, ———, and ———. 2007b. Accruals quality and internal control over financial reporting. *The Accounting Review* 82: 1141–1170.

Easley, D., and M. O'Hara. 2004. Information and the cost of capital. *Journal of Finance* 59: 1553–1583.

Easton, P. 2004. PE ratios, PEG ratios, and estimating the implied expected rate of return on equity capital. *The Accounting Review* 79: 73–95.

———, and S. Monahan. 2005. An evaluation of accounting-based measures or expected returns. *The Accounting Review* 80: 501–538.

Ecker, F., J. Francis, I. Kim, P. Olsson, and K. Schipper. 2006. A returns-based representation of earnings quality. *The Accounting Review* 81: 749–780.

Fama, E., and K. French. 1992. The cross-section of expected stock returns. *Journal of Finance* 46: 1575–1617.

Francis, J., R. LaFond, P. Olsson, and K. Schipper. 2004. Costs of equity and earnings attributes. *The Accounting Review* 79: 967–1010.

———, ———, ———, and ———. 2005. The market pricing of accruals quality. *Journal of Accounting and Economics* 39: 295–327.

Ge, W., and S. McVay. 2005. The disclosure of material weaknesses in internal control after the Sarbanes-Oxley Act. *Accounting Horizons* 19 (3): 137–158.

Gebhardt, W., C. Lee, and B. Swaminathan. 2001. Toward an implied cost of equity. *Journal of Accounting Research* 39: 135–176.

Gode, D., and P. Mohanram. 2003. Inferring the cost of equity using the Ohlson-Juettner model. *Review of Accounting Studies* 8: 399–431.

Guay, W., S. Kothari, and S. Shu. 2006. Properties of the implied cost of capital using analysts' forecasts. Working paper, Massachusetts Institute of Technology.

Hail, L., and C. Leuz. 2006. International differences in the cost of equity capital: Do legal institutions and securities regulation matter? *Journal of Accounting Research* 44: 485–531.

Hammersley, J., L. Myers, and C. Shakespeare. 2008. Market reactions to the disclosure of internal control weaknesses and to the characteristics of those weaknesses under Section 302 of the Sarbanes-Oxley Act of 2002. *Review of Accounting Studies* 13 (1): 141–165.

Hribar, P., and N. T. Jenkins. 2004. The effect of accounting restatements on earnings revisions and the estimated cost of capital. *Review of Accounting Studies* 9: 337–356.

Huber, P. J. 1967. The behavior of maximum likelihood estimates under non-standard conditions. *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, 221–233.

Jain, P., and Z. Rezaee. 2006. The Sarbanes-Oxley Act of 2002 and capital-market behavior: Early evidence. *Contemporary Accounting Research* 23: 629–654.

———, J. Kim, and Z. Rezaee. 2008. The Sarbanes-Oxley Act of 2002 and market liquidity. *Financial Review* (forthcoming).

Kasznik, R. 2004. Discussion of: The effect of accounting restatements on earnings revisions and the estimated cost of capital. *Review of Accounting Studies* 9: 357–367.

Lambert, R., C. Leuz, and R. Verrecchia. 2007a. Accounting information, disclosure, and the cost of capital. *Journal of Accounting Research* 45 (2): 385–420.

———, ———, and ———. 2007b. *Information Asymmetry, Information Precision, and the Cost of Capital.* Available at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=936429.

Li, H., M. Pincus, and S. Rego. 2008. Market reaction to events surrounding the Sarbanes-Oxley Act of 2002 and earnings management. *Journal of Law & Economics* (forthcoming).

Matsumoto, D. 2002. Management's incentives to avoid negative earnings surprises. *The Accounting Review* 77 (3): 483–514.

Merton, R. C. 1987. A simple model of capital market equilibrium with incomplete information. *Journal of Finance* 42: 483–510.

Michaels, A. 2003. Costs rise as U.S. businesses act to meet governance laws. U.S. edition. *Financial Times* (April): 15.

Ogneva, M., K. Raghunandan, and K. R. Subramanyam. 2007. Internal control weakness and cost of equity: Evidence from SOX Section 404 certifications. *The Accounting Review* 82: 1255–1297.

Ou, J., and S. Penman. 1989. Financial statement analysis and the prediction of stock returns. *Journal of Accounting and Economics* 11: 295–329.

Palmrose, Z-V., V. Richardson, and S. Scholz. 2004. Determinants of market reactions to restatement announcements. *Journal of Accounting and Economics* 37: 59–89.

Ribstein, L. E. 2002. Market vs. regulatory responses to corporate fraud: A critique of the Sarbanes-Oxley Act of 2002. *Journal of Corporation Law* 28 (1).

Romano, R. 2005. The Sarbanes-Oxley Act and the making of quack corporate governance. *Yale Law Journal* (June).

Securities and Exchange Commission (SEC). 2006. *Final Report of the Advisory Committee on Smaller Public Companies to the U.S. Securities and Exchange Commission.* Washington, D.C.: SEC.

Sloan, R. G. 1996. Do stock prices fully reflect information in accruals and cash flows about future earnings? *The Accounting Review* 71: 289–315.

Teoh, S., and T. Wong. 1993. Perceived auditor quality and the earnings response coefficient. *The Accounting Review* 68: 346–366.

Whisenant, J. S., S. Sankaraguruswamy, and K. Raghunandan. 2003. Market reactions to disclosure of reportable events. *Auditing: A Journal of Practice & Theory*: 181–294.

THE ACCOUNTING REVIEW
Vol. 83, No. 3
2008
pp. 705–734

Stock Options, R&D, and the R&D Tax Credit

Jennifer L. Brown
Arizona State University

Linda K. Krull
The University of Texas at Austin

ABSTRACT: Two separate streams of research find evidence that firms decrease R&D spending to meet earnings benchmarks and that the R&D tax credit increases R&D spending. However, these studies do not consider stock option exercises by R&D employees that likely influence R&D spending decisions because they increase R&D tax credits without reducing reported earnings. This study extends both areas of research by incorporating R&D tax credits from stock option exercises into the R&D spending decision. We find evidence that firms reduce R&D spending by 0.46 percent of total assets to avoid earnings decreases, but R&D tax credits generated by stock option exercises offset this decrease by 16 percent on average and up to 42 percent for a fully taxable firm. The results of this study suggest that the tax benefits of R&D-related stock option exercises are important considerations in studies that investigate myopic R&D investment and in studies that investigate the effect of the R&D tax credit on R&D spending.

Keywords: *research and development; tax credits; stock options; earnings management.*

Data Availability: *All data is publicly available from sources identified in the text.*

I. INTRODUCTION

Recent research finds evidence that while accruals management decreases following passage of the Sarbanes-Oxley Act, real earnings management increases during this time period (Cohen et al. 2008). Though real earnings management has been studied at length in the accounting literature (see e.g., Baber et al. 1991; Bushee 1998; Gunny 2005; Zang 2005; Roychowdhury 2006; Cohen et al. 2008), this recent shift has increased concern over real earnings management, highlighting the importance of understanding factors that affect its use to meet earnings benchmarks.[1]

---

[1] See Healy and Wahlen (1999) and Fields et al. (2001) for reviews of this research.

We thank Dan Dhaliwal, Jeffrey Hales, Lisa Koonce, Steve Limberg, Matthew Magilke, Bill Mayew, Lillian Mills, John Robinson, Andy Schmidt, Connie Weaver, two anonymous referees, and workshop participants at The University of Texas, University of Utah, the 2006 Lone Star Accounting Research Conference at Texas A&M University, and the 2006 American Accounting Association Annual Meeting for their helpful comments.

Editor's note: This paper was accepted by Dan Dhaliwal.

Submitted December 2005
Accepted November 2007

One commonly cited example of real earnings management is myopic research and development (hereafter, R&D) investment, which occurs when managers forgo positive net present value projects to meet earnings benchmarks. Empirical and anecdotal evidence indicates that managers cut R&D to meet earnings benchmarks in spite of potential long-run detriments to investors (Baber et al. 1991; Bushee 1998; Roychowdhury 2006). Existing studies also find evidence that corporate governance mechanisms, such as institutional investors (Bushee 1998) and compensation committees (Cheng 2004), mitigate managers' opportunistic reductions in R&D. We investigate an additional factor that potentially mitigates myopic R&D investment in the presence of short-term earnings constraints: R&D tax credits from stock option exercises.

The effect of R&D tax credits from stock option exercises on myopic R&D investment is important to understand because the firms that benefit from R&D tax credits are often the same firms that use stock options extensively to compensate employees; research shows that more R&D-intensive firms use more stock options to compensate employees (Ryan and Wiggins 2002). Moreover, R&D tax credits generated by stock option exercises decrease the financial reporting costs of R&D spending. Therefore, stock option exercises can be an important factor in R&D spending decisions for many R&D intensive firms, and omitting this important factor can result in over- or underestimating myopic R&D investment for some firms. Further, understanding the effect of R&D tax credits from stock option exercises on myopic R&D investment can help investors and researchers identify firms that are more or less likely to engage in this type of real earnings management.

R&D tax credits encourage managers to invest in more R&D by decreasing the after-tax cost of R&D spending. Specifically, the R&D tax credit allows firms to reduce their federal tax liability by 20 percent of qualified R&D expenditures in excess of a base amount (hereafter we refer to R&D in excess of the base amount as "excess R&D"), but limits "excess R&D" to 50 percent of qualified R&D.[2,3] Qualified R&D expenditures include in-house payments for salaries and wages, supplies, and 65 percent of contract R&D payments.

A potentially large component of qualified R&D expenditures occurs when R&D employees exercise stock options. Specifically, qualified expenditures for salaries and wages include the tax deduction from R&D-related stock option exercises (hereafter, R&D option exercises), equal to the spread between the market value and strike price of stock options exercised by employees conducting R&D. This component of qualified R&D expenditures is arguably economically significant. Our sample estimates of R&D option exercises range from $0 to $5.8 billion. The mean is $16.21 million, and on average more than triples excess R&D, which, in turn, more than triples total R&D tax credits.

R&D option exercises are important in the context of myopic R&D investment because they decrease the financial reporting costs of R&D spending. The earnings effect of R&D option exercises appears minimal at first glance because, during the pre-FAS No. 123R period we study, firms generally credit the tax benefits of stock option exercises to additional paid-in-capital, which does not reduce income tax expense on the financial statements.

---

[2] The base amount equals the maximum of the fixed base percentage times the average of gross receipts over years $t-4$ through $t-1$ and 50 percent of current-year qualified R&D. The fixed base percentage equals the minimum of 16 percent and the sum of qualified R&D from 1984–1988 divided by the sum of gross receipts over the same period. If a firm has fewer than three years of both gross receipts and qualified R&D from 1984 to 1988, then the fixed base percentage equals 3 percent.

[3] In addition to taking the credit, firms can deduct qualified R&D net of the credit from taxable income. For example, assume a firm has $100 of qualified R&D expenditures in excess of the base amount. Using the current credit rate of 0.20, the firm can claim a credit of 0.20*100 = $20. It can deduct S100 − $20 = $80 from taxable income. Assuming the firm's tax rate is 35 percent, the tax benefit of the excess R&D is 20 + (80*.35) = $48. Without the R&D credit the tax benefits of the excess R&D would be $100*.35 = $35.

However, as we demonstrate in the next section, R&D option exercises increase the amount of non-option R&D in excess of the base amount and, therefore, increase the amount of R&D tax credits reported in earnings. Because R&D option exercises themselves do not directly affect earnings while the credits they generate reduce income tax expense, R&D option exercises reduce the after-tax earnings effect of R&D.

As a result of this decrease in financial reporting costs, as R&D option exercises increase, firms near an earnings benchmark can invest in more R&D before falling short of their earnings goal. This financial reporting consequence suggests that the results found in prior studies of myopic R&D investment vary with the magnitude of firms' tax credits from R&D option exercises: as tax credits from R&D option exercises increase, firms are less likely to forgo or postpone R&D spending to meet earnings benchmarks.[4] However, prior research ignores the stock option component of R&D spending. Therefore, we investigate whether tax credits from R&D option exercises mitigate myopic R&D investment related to short-term earnings constraints.

We examine the mitigating role of R&D option exercises by studying how short-term earnings constraints and R&D option exercises affect the sensitivity of firms' R&D spending decisions to the R&D tax credit. Our sample includes 2,137 firm-years from 1992 through 2002. Within this sample, we define firm-years with short-term earnings constraints as those facing a potential earnings decrease, that is those for which maintaining the level of R&D spending in the previous year would prevent them from reporting increasing earnings (Baber et al. 1991; Bushee 1998).

Like existing research, we find evidence that R&D spending is positively related to the R&D credit rate, and that firms with potential earnings decreases reduce R&D relative to other firms. Extending this research, we also find that R&D option exercises increase firms' sensitivity to the credit rate, i.e., the positive association between the tax credit rate and R&D spending is increasing in R&D option exercises. Consistent with a mitigating role of tax credits from R&D option exercises, we find that the effect of R&D option exercises on firms' sensitivity to the credit rate is even higher for firms with potential earnings decreases. This result suggests that tax credits from R&D option exercises reduce the amount by which firms decrease R&D during years with stronger incentives to increase earnings. Specifically, we find evidence that firms reduce R&D spending by 0.46 percent of total assets to avoid earnings decreases, but R&D tax credits generated by stock option exercises offset this decrease by 16 percent on average and up to 42 percent for a fully taxable firm.

This study makes three contributions to the literature. First, we extend the literature that investigates real earnings management (Baber et al. 1991; Bushee 1998; Gunny 2005; Zang 2005; Roychowdhury 2006; Cohen et al. 2008; among others) by investigating whether R&D tax credits and R&D option exercises mitigate myopic R&D investment. The evidence in this study suggests that R&D tax credits generated by R&D option exercises significantly reduce the magnitude of R&D cuts to avoid earnings decreases. This evidence contributes to our understanding of factors that influence investment and financial reporting decisions and suggests that, in studies of myopic R&D investment, both stock option exercises and R&D tax credits are important considerations.

Second, this study contributes to the literature that investigates the effect of income taxes on earnings management. Existing tax literature finds evidence that firms use income tax accruals to manage earnings (see e.g., Phillips et al. 2003; Dhaliwal et al. 2004) but does not investigate income taxes in the context of real earnings management. We find evidence that taxes play a mitigating role in one form of real earnings management, myopic R&D investment. Thus, the results of this study are useful for future research that can identify other circumstances in which income tax laws and financial reporting standards for income taxes mitigate or exacerbate real earnings management.

---

[4] Similarly, but from another perspective, as tax credits from R&D option exercises increase, the increase in earnings from cutting a given amount R&D becomes smaller. This effect increases the size of R&D cuts needed to meet the earnings goal. Therefore, myopic R&D investment becomes more costly relative to other methods of real earnings management and accruals management, making firms less likely to cut R&D.

Third, this study contributes to the literature that investigates the effectiveness of R&D tax incentives. We provide some evidence that R&D option exercises increase the sensitivity of R&D spending to the R&D credit rate. In particular, the evidence in this study suggests that stock options can be an important component of the R&D tax credit for firms with large amounts of R&D option exercises. Accordingly, this study provides useful information for future research investigating the effectiveness of the credit.

One potential limitation of this study lies in the timing of option grants and the exercise of those options. If management grants stock options to R&D employees, then the tax deduction generated when employees exercise those options occurs several years later. Therefore, it is not management's decision to increase R&D option exercises. Though option exercises are not directly in management's control, the tax benefits of those exercises can affect management's tax and financial planning. Conversations with partners from the R&D tax practices of two large national firms confirm that since employee option exercises comprise a significant portion of a firm's R&D tax credit, managers and their tax advisors routinely estimate likely employee option exercises and incorporate these estimates into their tax planning. In addition, empirical evidence suggests that managers consider the tax benefits from option exercises when choosing their capital structure (Graham et al. 2004).

A second limitation of this study is that we must estimate management's expectation of R&D option exercises. To reduce the possibility that our results are driven by assumptions we make in estimating management's expectation of R&D option exercises, we replicate our results using several estimates of R&D option exercises. We also perform sensitivity tests to control for measurement error caused by differences between actual option exercises and anticipated exercises. However, to the extent that our estimates of R&D option exercises differ from management's anticipated exercises and to the extent that we are unable to completely control for unexpected exercises, our empirical tests contain measurement error.

The paper proceeds as follows: Section II develops the hypotheses, Sections III discusses the research methodology, Section IV discusses the results, and Section V concludes.

II. HYPOTHESIS DEVELOPMENT

In this section, we develop predictions for how the tax and earnings effects of stock option exercises interact in the context of short-term earnings constraints to mitigate myopic R&D investment. First, we discuss our prediction for the effect of short-term earnings constraints on R&D spending based on the results of existing myopic R&D investment studies. Second, we develop our hypothesis for how tax credits from R&D option exercises mitigate myopic R&D investment in the presence of short-term earnings constraints following the model of R&D investment in Klassen et al. (2004).

Short-Term Earnings Constraints

Existing research finds evidence that firms decrease R&D to meet earnings benchmarks. Baber et al. (1991) find that firms decrease R&D to avoid earnings decreases. In addition, Roychowdhury (2006) finds evidence that firms attempt to avoid losses by reducing discretionary expenditures such as R&D and capital expenditures. Based on this existing evidence, we expect that firms with short-term earnings constraints reduce R&D relative to other firms. To investigate whether this result holds in our sample, we define firms with short-term earnings constraints as those facing a potential earnings decrease, that is, those for which maintaining the level of R&D spending in the previous year would prevent them from reporting increasing earnings, and test the following hypothesis:

H1: Firms with potential earnings decreases reduce R&D spending relative to other firms.

Short-Term Earnings Constraints and the R&D Tax Credit

Existing evidence suggests that R&D spending is increasing in the R&D credit rate. For example, Berger (1993) analyzes the R&D spending of 231 firms and finds that from 1982 to 1985, the R&D credit induced \$1.74 of R&D spending per dollar of forgone tax revenue.[5] Gupta et al. (2004) and Klassen et al. (2004) construct similar estimates and find that during the 1990s, the credit generated between \$2.40 and \$2.96 of R&D spending per dollar of forgone tax revenue.

To investigate the role of tax credits from R&D option exercises in mitigating myopic R&D investment, we build on this research by studying the effects of short-term earnings constraints and R&D option exercises on the sensitivity of R&D spending to the R&D credit. To develop our predictions, we follow the model of R&D investment in Klassen et al. (2004). In that paper, they develop an expression for the expected after-tax return on R&D investment assuming that firms without financial constraints invest in all projects with positive values for this expression and firms with financial constraints invest until they reach the constraint.[6] Then they derive an expression for the effect of the credit rate on R&D spending for financially constrained and unconstrained firms. We briefly summarize this analysis and then incorporate the effect of R&D option exercises.

Brief Explanation of Klassen et al. (2004)

From Klassen et al. (2004), the profitability index for R&D expenditures (η), equal to the ratio of the after-tax net cash flows to the after-tax investment, can be written as:

$$\eta = \frac{EPV(CF)(1-t) - RD[1-t-CR(1-t)]}{RD[1-t-CR(1-t)]} = \frac{1}{1-CR} \times \frac{EPV(CF)}{RD} - 1 \quad (1)$$

where $EPV(CF)$ equals the expected present value of before-tax cash flows, t equals the tax rate at which cash flows are taxable and R&D expenditures net of the R&D credit are deductible, RD equals qualified R&D expenditures, and CR equals the R&D credit rate.

---

[5] Other studies that find evidence that the R&D credit increases R&D spending during this time period include Swenson (1992), McCutchen (1993), and Hall (1993). However, results during this period provide mixed evidence on the effectiveness of the R&D tax credit because the moving average base amount originally used to calculate the credit created a disincentive to invest in R&D, i.e., in some situations the effective rate of the credit was negative (see e.g., Eisner et al. 1984; Altshuler 1988). For example, Tillinger (1991) estimates that from 1981 to 1985, the R&D credit induced \$460 million of R&D expenditures at a cost of \$2.4 billion, which translates into only \$0.19 of additional R&D per dollar of forgone tax revenue.

[6] Klassen et al. (2004) mention two types of financial constraints, financial reporting constraints, and capital constraints. Building on existing real earnings management literature, our study focuses on financial reporting constraints and, more specifically, short-term earnings constraints.

Assuming firms without financial constraints invest in all R&D projects with positive net present values, Klassen et al. (2004) set Equation (1) equal to 0 and rearrange to isolate the effect of the credit rate, which results in the following expression:

$$\frac{EPV(CF)}{RD} = 1 - CR. \tag{2}$$

Equation (2) suggests that when firms have no financial constraints R&D spending is increasing in the credit rate.

For a firm with financial constraints, Klassen et al. (2004) show that the decision to invest in R&D changes in two ways. First, instead of investing in all projects with a positive net present value, the firm invests until it reaches the constraint. Second, since the tax savings on all R&D investment affects how much the firm can invest before it reaches the constraint, the proportion of R&D that exceeds the base amount, i.e., the ratio of excess R&D to total R&D, becomes relevant to the investment decision. For a firm with financial constraints, the after-tax investment in R&D must be less than the constraint. Therefore, the investment problem becomes:

$$\begin{aligned} &\max \frac{1}{1 - \theta CR} \times \frac{EPV(CF)}{RD} - 1 \\ &\text{s.t. } RD(1 - t) - [RD \times \theta CR(1 - t)] \leq C \end{aligned} \tag{3}$$

where C represents the financial constraint and θ equals the proportion of R&D that exceeds the base amount. Setting the constraint to an equality and writing the constraint as a fraction of RD, the condition becomes:

$$\frac{C}{RD} = 1 - t - \frac{RD \times \theta CR(1 - t)}{RD}. \tag{4}$$

This expression suggests that the effect of a change in the credit rate on $\frac{C}{RD}$ is:

$$\frac{\partial C/RD}{\partial CR} = -\theta(1 - t). \tag{5}$$

Equation (5) suggests that, when firms face financial constraints, R&D spending is increasing in the credit rate, and the magnitude of this effect is increasing in θ.

In summary, Klassen et al. (2004) show that when firms have no financial constraints, R&D spending is increasing in the R&D credit rate. When firms face financial constraints, R&D spending is increasing in the credit rate times the proportion of R&D that exceeds the base amount.

Incorporation of R&D Option Exercises

We extend Klassen et al. (2004) to incorporate R&D option exercises into the R&D investment decision by writing the statutory definition of the proportion of R&D that exceeds the base amount θ_s, as follows:

$$\theta_s = \frac{RD + SO - FB}{RD + SO} \tag{6}$$

where RD equals non-option R&D expenditures, SO equals R&D option exercises, FB equals the base amount, and $0 \leq \theta_s \leq 0.5$. Equation (6) is increasing in SO, which suggests that the proportion of R&D that exceeds the base amount is increasing in R&D option exercises.

In this study we focus on one type of financial constraint, short-term earnings constraints. Firms with short-term earnings constraints have incentives to manage earnings. That is, firms near an earnings benchmark have incentives to increase earnings to avoid missing the earnings target. In this context, the importance of θ lies in its effect on reported earnings. Therefore, our interest in this study is the proportion of non-option R&D that exceeds the base amount because only non-option R&D and credits on non-option R&D affect reported earnings during the period we study. To focus on the proportion of non-option R&D that exceeds the base amount, we write θ as follows:

$$\theta = \frac{RD + SO - FB}{RD}. \tag{7}$$

Replacing θ in Equation (5) with Equation (7) shows that when firms have short-term earnings constraints, the sensitivity of R&D investment to the credit rate is increasing in R&D option exercises. Intuitively, as the credit rate increases, a firm near an earnings benchmark can invest in more R&D before reaching the constraint, and the effect of an increase in the credit rate is larger as R&D option exercises increase because the credit rate applies to a larger proportion of non-option R&D. Since firms can invest in more R&D before reaching the constraint, we expect that they are less likely to forgo or postpone positive net present value projects to meet an earnings benchmark. In other words, we expect that R&D tax credits generated by R&D option exercises decrease the financial reporting costs of non-option R&D spending and therefore decrease firms' opportunistic reductions in R&D to meet earnings benchmarks.[7]

Further, Equation (2) suggests that when firms have no financial constraints the sensitivity of R&D investment to the credit rate is unrelated to θ and, therefore, is also unrelated to R&D option exercises. We exploit this difference in the effect of R&D option exercises to investigate whether R&D tax credits generated by R&D option exercises mitigate myopic R&D investment. Specifically, we test whether the effect of R&D option exercises on the sensitivity of firms' R&D spending decisions to the R&D credit rate is greater for firms identified as likely to have short-term earnings constraints, which we define as firms with potential earnings decreases, than for the remainder of the sample. Such evidence would suggest that firms with more tax credits from R&D option exercises reduce R&D less for earnings objectives.

Note, however, that increasing θ only helps a firm meet its earnings benchmark if the increase results in an income tax benefit on the financial statements. Although firms can carry forward unused credits, if a firm has large or persistent tax losses, then it is more likely to record a deferred tax asset valuation allowance and less likely to record a tax benefit when R&D tax credits increase.[8] In addition, the tax law specifies that θ_s cannot exceed 50 percent. For firms with θ_s equal to 0.50, the effect of R&D option exercises on excess non-option R&D is reduced or eliminated, making the proportion of R&D eligible for the credit less sensitive to R&D option exercises. To increase the power of our tests, we focus on firms for which earnings are likely to be sensitive to tax credits generated by R&D option exercises. We label these firms credit-sensitive firms, which we define as those that are below the maximum θ_s before considering option exercises and that are not reporting persistent tax losses. Therefore, we test the following hypothesis:

---

[7] Because during our sample period the tax benefits of SO are credited to additional paid-in capital and thus do not affect earnings, the effect of SO on earnings depends on how the firm allocates excess R&D between RD and SO. Equation (7) implies that excess R&D is allocated first to RD and then to SO. Equation (6) implies that the allocation is proportional. For both allocations, the benefit of the credit is increasing in SO. However, the effect is larger using the allocation implied by Equation (7). Tax law limits Equation (6) to be between 0 and 0.5. Financial reporting for the tax benefits of stock options limit Equation (7) to be between 0 and 1. Financial reporting standards do not specify the procedure with which to make this allocation.

> **H2:** Among credit-sensitive firms, the effect of R&D option exercises on the relation between the R&D credit rate and non-option R&D spending is greater for firms with potential earnings decreases than for other firms.

III. RESEARCH DESIGN

Empirical Model

To test our first hypothesis, we estimate the following ordinary least-squares regression of changes in non-option R&D spending:

$$\begin{aligned}\Delta RD_{i,t} = {} & \alpha_0 + \alpha_1 EarnDec_{i,t} + \alpha_2 Credit_{i,t} + \alpha_3 RDoptions_{i,t} \\ & + \alpha_4 RDoptions{*}Credit_{i,t} + \alpha_5 TaxRate_{i,t} + \alpha_6 Q_{i,t} \\ & + \alpha_7 CFshort_{i,t} + \alpha_8 IndRD_{i,t} + \alpha_9 GNP_t + \alpha_{10} ln(Sales)_{i,t} \\ & + \alpha_{11} Grants_{i,t} + \alpha_{12} Repurchases_{i,t} + \alpha_{13} ROA_{i,t} + \varepsilon_{i,t}. \qquad (8)\end{aligned}$$

See Figure 1 for variable definitions.

The dependent variable is the change in R&D from year $t-1$ to year t.[9] The coefficient on *EarnDec*, equal to 1 for firm-years with potential earnings decreases, and 0 otherwise, tests H1. We define firm-years with potential earnings decreases as firm-years for which year t earnings before extraordinary items and R&D is below target earnings by an amount within year $t-1$ R&D (Baber et al. 1991; Bushee 1998).[10] Target earnings equals year $t-1$ earnings before extraordinary items and R&D. Hypothesis 1 predicts that firms with potential earnings decreases reduce R&D relative to other firms. Therefore, we predict that α_1 is negative.

Credit measures the present value of R&D tax credits generated by the last dollar of R&D spending. The R&D credit rate is 20 percent of excess R&D. Therefore, if R&D is less than the base amount, then we set *Credit* equal to 0. If a firm's current tax liability

---

[8] Statement of Financial Accounting Standards No. 109 prescribes that firms record a deferred tax asset for tax credit carryforwards in the same manner they would for the future tax benefits of deductible temporary differences and net operating loss carryforwards. The firm records a deferred tax asset valuation allowance if it deems that it is more likely than not than some or all of the deferred tax asset will not be realized.

[9] Studies of the R&D tax credit typically estimate an empirical model with year t R&D as the dependent variable and year $t-1$ R&D as an explanatory variable. These studies include year $t-1$ R&D on the right-hand side to control for the tendency of R&D projects to require multi-year commitments. This specification is econometrically similar to ours but allows the coefficient on year $t-1$ R&D to be different from 1. When we estimate this specification, the coefficient on year $t-1$ R&D is very close to 1. To avoid a unit-root problem, we use the change in R&D as the dependent variable. All inferences are unchanged if we estimate the model with year t R&D as the dependent variable and year $t-1$ R&D as an explanatory variable.

[10] Baber et al. (1991) and Bushee (1998) define target earnings as earnings before taxes and R&D. We use earnings before R&D and extraordinary items and after taxes because we are interested in the tax effects of R&D.

FIGURE 1
Variable Definitions

Variable Name	Definition (Compustat Data Item Numbers in Parentheses)	Predicted Sign
ΔRD	Year *t* R&D expense reported in the financial statements (data46) minus year *t*−1 R&D expense, divided by total assets at the beginning of year *t* (data6).	
EarnDec	An indicator variable equal to 1 for firm-years for which earnings before extraordinary items and R&D (data18 + data46) is below target earnings by an amount within year *t*−1 R&D. Target earnings equals year *t*–1 earnings before extraordinary items and R&D.	(−)
Credit	The R&D tax credit rate equal to *PVCR**(1 − *TaxRate*) where *PVCR* equals the present value of the R&D credit rate for firms with non-option R&D (data46) plus ***R&D Options*** greater than *Fixed Base*, and 0 otherwise.	(+)
Fixed Base	The maximum of 0.50*(R&D + *R&D Options*) and the fixed base percentage times mean sales (data12) from *t*−4 through *t*−1. The fixed base percentage equals the minimum of 16 percent and the mean ratio of R&D to sales from 1984–1988. If data are unavailable to calculate the mean ratio of R&D to sales from 1984 to 1988, then the fixed base percentage equals 3 percent.	
RDoptions	R&D option exercises, ***R&D Options***, divided by beginning of year total assets (data6). ***R&D Options*** equals the value realized from options exercised by non-executives times the average ratio of R&D (data46) to operating expenses (data12–data178) from 1987–2002.	(+/−)
TaxRate	0.35 if the firm has (1) positive taxable income (*TI*) in year *t* and (2) no net operating loss (NOL) carry-forward from year *t*−1, 0.175 if the firm has (1) or (2), and 0 if the firm has both (1) and (2).	(+)
Taxable Income (TI)	Pre-tax book income (data170) minus (deferred tax expense from the statement of cash flows [(data126)/0.35]) minus (the tax deduction from executive option exercises divided by the ratio of the number of options granted to executives to the total number of options granted for the year).	
Q	Tobin's q, measured as [market value of equity (data199*data25) + short-term debt (data34) + long-term debt (data9) + the book value of preferred stock (data130)] scaled by total assets (data6).	(+)
CFShort	Cash flow shortfall before R&D measured as [common (data21) and preferred (data19) dividends + cash flow from investing (data311) − cash flow from operations (data308) − R&D (data46)] scaled by beginning of year total assets (data6).	(−)
IndRD	The mean ratio of R&D (data46) to total assets at the beginning of the year (data6) for all firms in firm *i*'s industry based on Fama and French (1997) industry classifications.	(+)
GNP	The log of year *t* real gross national product.	(+)
ln(Sales)	Firm size measured as the natural logarithm of sales (data12).	(+/−)

(*continued on next page*)

FIGURE 1 (continued)

Variable Name	Definition (Compustat Data Item Numbers in Parentheses)	Predicted Sign
Grants	The Black-Scholes value of [option grants to executives divided by the percentage of total option grants made to executives (from ExecuComp)] divided by total assets at the beginning of the year (data6).	(+/−)
Repurchases	Share repurchases, measured as purchases of common and preferred stock (data115) − the change from year $t-1$ to year t in preferred stock (data130 − data175), divided by total assets at the beginning of the year (data6).	(−)
ROA	Pre-R&D return on assets, measured as net income before R&D (data172 + data46) divided by lagged total assets (data6).	(+)

is less than total R&D credits, then it can carry unused credits back one year and forward 20 years. Therefore, for firms with excess R&D and a tax liability in the current or previous year greater than total R&D credits, *Credit* equals 0.20. For firms whose tax liability in the current and previous year is less than total R&D credits, we incorporate the ability to carry unused credits forward into our measure of the credit rate by forecasting taxable income through the 20-year carry forward period. We use the forecasted stream of taxable income to determine the period during which the firm will benefit from the credit, and we discount the credit rate to the current period. This process results in our estimate of the present value of the credit rate (*PVCR*).[11] From Equation (5) the effect of a change in the credit rate on R&D is $\theta(1 - t)$. To isolate the effect of R&D option exercises on R&D, we define *Credit* as *PVCR**(1 − *TaxRate*). Consistent with existing literature, we expect current-year R&D expenditures to be positively related to the tax benefit of R&D expenditures. Therefore, we predict that α_2 is positive.

RDoptions measures the tax deduction from stock options exercised by employees engaged in qualified research (which we refer to as R&D option exercises).[12] Employees are engaged in qualified research if they conduct research or directly supervise research, but direct supervision does not include higher-level managers. Therefore, we estimate R&D option exercises, *R&D Options*, using the tax deduction from non-executive employee option exercises from ExecuComp as follows:

R&D Options = *Employee Exercises**R&D Ratio, and
Employee Exercises = [(Executive Exercises/Executive Percent)*(1 − Executive Percent)],

where Executive Exercises equals the value realized from options exercised by executives, R&D Ratio equals the firm-specific ratio of R&D to total operating expenses averaged from 1987 through 2002, and Executive Percent equals the ratio of the number of options granted to executives to the total number of options granted during the year. This estimate assumes that non-executives have the same weighted average exercise price as executives during the year. *RDoptions* equals *R&D Options* divided by beginning of year total assets.[13]

---

[11] See the Appendix for a more detailed discussion of the calculation of *PVCR*.

[12] Tax law distinguishes between two types of stock options. Nonqualified stock options result in a corporate deduction for the spread between the stock price and exercise price at the date of exercise and income in the same amount to the employee. Incentive stock options result in no deduction to the corporation and no income to the employee at the date of exercise. As in Graham et al. (2004), we make the simplifying assumption that all reported options are nonqualified.

Ryan and Wiggins (2002) find that R&D is increasing in stock option grants to employees. To the extent that firms that grant more stock options to R&D employees also have higher tax deductions from stock option exercises, we expect a positive relation between ΔRD and *RDoptions*. However, Bens et al. (2002) predict that firms decrease R&D to fund share repurchases, which offset the earnings per share dilution of option exercises. Consistent with this expectation, they find a negative relation between R&D and executive option exercises. Therefore, we make no directional prediction for α_3.

*RDoptions*Credit* captures the effect of R&D option exercises on the sensitivity of R&D spending to the credit rate. Our theory predicts that R&D option exercises increase the positive relation between R&D and the credit rate, which suggests a positive coefficient on *RDoptions*Credit*. However, we expect this relation only for credit-sensitive firm-years that face a potential earnings decrease. Therefore, we make no prediction for α_4 when we estimate Equation (8) for the full sample.

Following Berger (1993), Gupta et al. (2004), and Klassen et al. (2004), Equation (8) includes variables that control for other factors that influence R&D. Firms with higher marginal tax rates get more benefit from deducting R&D expense. To control for this effect, we include a trichotomous variable (*TaxRate*) that captures the firm's current-year tax rate after R&D and option exercises and expect a positive coefficient on this variable (Shevlin 1990). We include Tobin's q (*Q*), to control for the firm's investment opportunity set (Berger 1993). We expect that firms with more investment opportunities invest in more R&D and, therefore, expect a positive coefficient on *Q*.

Myers and Majluf (1984) suggest that firms prefer internal financing to external financing. As a firm's internal funds available for R&D increase, the cost of financing R&D decreases. We control for the availability of internal funds by including an estimate of the firm's cash flow shortfall (*CFshort*) similar to Core and Guay (2001). Firms with more cash flow shortfall have less cash available for R&D expenditures; thus, we expect a negative coefficient on *CFshort*. We include the mean ratio of R&D to total assets at the beginning of the year for the firm's industry, based on the Fama and French (1997) industry classifications (*IndRD*) and real gross national product (*GNP*) to control for macroeconomic factors that influence R&D spending. We expect changes in R&D to be positively related to the mean R&D for the firm's industry and, therefore, expect a positive coefficient on *IndRD*. *GNP* is increasing in overall growth in the U.S. economy. We expect firms to invest in more R&D during years with more overall economic growth. Therefore, we expect a positive coefficient on *GNP*. We include *ln*(*Sales*) as a control for size. Consistent with existing studies, we make no prediction for this variable (Bens et al. 2002; Gupta et al. 2004).

Existing studies find evidence that firms with more R&D intensity use more options to compensate management and employees (Ryan and Wiggins 2002; Cheng 2004). In addition, Core and Guay (2001) find that stock option grants are a significant determinant of stock option exercises. Because option grants are positively related to both our dependent variable and the independent variable of interest, we include the Black-Scholes value of stock option grants divided by total assets at the beginning of the year, *Grants*, to control for the potential bias created by correlated omitted variables. Existing evidence suggests a positive relation between option grants and R&D. Conversely, if a firm faces financial constraints, then it has an incentive to replace cash compensation with option grants, and can replace cash compensation for R&D employees with options grants. This substitution effect suggests a negative relation between option grants and R&D. Therefore, we make no prediction for the coefficient on *Grants*.[14]

---

[13] We check the reasonableness of this measure of R&D option exercises by comparing it to a measure calculated using footnote data. The correlation between the footnote and ExecuComp measures for 416 firm-year observations with footnote data available is 0.56, suggesting a reasonable amount of accuracy for our ExecuComp measure. We discuss the footnote measure in more detail in the "Sensitivity Tests" subsection.

Existing evidence also suggests that firms increase share repurchases to offset the earnings per share dilution of option exercises, which decreases the capital available for R&D spending (Bens et al. 2002). We control for this potential effect of share repurchases on R&D spending by including *Repurchases* as an explanatory variable, and we expect a negative coefficient on this variable. We also include *ROA*, a measure of profitability before R&D, to control for the possibility that firms both invest in more R&D and have more option exercises during more profitable years. Thus, we expect the coefficient on *ROA* to be positive.

We investigate whether R&D option exercises mitigate myopic R&D investment by adding interactions of *EarnDec* with *Credit*, *RDoptions*, and *RDoptions*Credit* and estimating the following regression for credit-sensitive firm-years:

$$\begin{aligned}\Delta RD_{i,t} = {} & \gamma_0 + \gamma_1 EarnDec_{i,t} + \gamma_2 Credit_{i,t} + \gamma_3 RDoptions_{i,t} \\ & + \gamma_4 RDoptions{*}Credit_{i,t} + \gamma_5 EarnDec{*}Credit_{i,t} \\ & + \gamma_6 EarnDec{*}RDoptions_{i,t} + \gamma_7 EarnDec{*}RDoptions{*}Credit_{i,t} \\ & + \gamma_8 TaxRate_{i,t} + \gamma_9 Q_{i,t} + \gamma_{10} CFshort_{i,t} + \gamma_{11} IndRD_{i,t} + \gamma_{12} GNP_t \\ & + \gamma_{13} ln(Sales)_{i,t} + \gamma_{14} Grants_{i,t} + \gamma_{15} Repurchases_{i,t} + \gamma_{16} ROA_{i,t} + \eta_{i,t}. \quad (9)\end{aligned}$$

See Figure 1 for variable definitions.

As in Equation (8), we expect the coefficient on *EarnDec* (γ_1) to be negative, and we make no prediction for the coefficient on *RDoptions*Credit* (γ_4). We expect that tax credits from R&D option exercises mitigate the negative effect of short-term earnings constraints on R&D. Consistent with this expectation, H2 predicts that the effect of R&D option exercises on the positive relation between *Credit* and ΔRD is significantly greater for firm-years with potential earnings decreases than for the remainder of the sample. The coefficient on *EarnDec*RDoptions*Credit* (γ_7) tests this prediction. Therefore, we expect γ_7 to be positive.

*EarnDec*Credit* controls for the change in the effect of *EarnDec* on ΔRD as *Credit* increases. R&D tax credits from non-option R&D decrease the after-tax earnings effect of

---

[14] These relations also suggest a potential problem with endogeneity. To test for endogeneity, we conduct a Hausman test and find that endogeneity is not a problem in our empirical tests ($t = -0.17$; $p = 0.8656$). Nevertheless, we re-estimate the tests in Table 5 using a two-stage least squares analysis with ΔRD, *RDoptions*, and *RDoptions*Credit* as endogenous variables. In the two-stage least squares test, we estimate the model separately for firms with potential earnings decreases (*EarnDec* = 1) and other firms (*EarnDec* = 0). We define *RDoptions* and *RDoptions*Credit* as a function of *Block* (the percent of shares held by institutions that own 5 percent or more), *Leverage* (the ratio of total debt to total assets), *AccNoise* (the variance in the annual change in ROE divided by the variance in the annual stock return over the sample period as in Yermack [1995]), *Emp* (the number of employees divided by total assets at the beginning of the year), *3yrReturn* (stock return from the beginning of year $t-2$ to the end of year t), *Phigh* (the number of months during the previous 12 months in which the closing stock price reached a 12-month high), *Plow* (the number of months during the previous 12 months in which the closing stock price reached a 12-month low), *Inst* (the percent of shares outstanding owned by institutions), *Grants*, and *Repurchases*. Inferences are similar to those reported in Table 5. The coefficient on *RDoptions*Credit* is 175.79 ($t = 1.84$) for firms with potential earnings decreases and -63.90 ($t = -1.52$) for other firms.

R&D, which suggests that *Credit* mitigates the negative effect of potential earnings decreases on ΔRD. Thus, we would expect a positive coefficient on *EarnDec*Credit.* However, *EarnDec*Credit* also controls for the difference in the effect of *Credit* on ΔRD for firm-years with *EarnDec* equal to 1, relative to firm-years with *EarnDec* equal to 0. Equation (2) suggests that the change in R&D for a change in the credit rate is one-to-one for firms without potential earnings decreases (*EarnDec* = 0), while Equation (5) suggests that the change in R&D for a change in the credit rate is $\theta(1 - t)$ for firms with potential earnings decreases (*EarnDec* = 1). Since $\theta(1 - t)$ is less than 1 (recall that θ is the proportion of R&D that exceeds the base amount and t is the tax rate), we would expect the coefficient on *Credit* to be lower for firm-years with potential earnings decreases suggesting a negative coefficient on *EarnDec*Credit.* Because these two effects are offsetting we make no prediction for the coefficient on *EarnDec*Credit* (γ_5).

*EarnDec*RDoptions* controls for the difference in the relation between *RDoptions* and ΔRD for firms with potential earnings decreases. Our theory predicts that the coefficient on this variable is not significantly different from zero since *RDoptions* only affects ΔRD as a multiple of *Credit,* and this effect is captured by *EarnDec*RDoptions*Credit.* However, Bens et al. (2002) find that R&D is decreasing in the value of option exercises and share repurchases are increasing in the value of option exercises. They interpret this result as evidence that firms decrease R&D to fund share repurchases undertaken to offset the dilutive effect of option exercises on earnings per share. Similarly, firms with potential earnings decreases may be more likely to decrease R&D to directly increase earnings and offset the dilutive effect of option exercises. Thus, we expect a negative coefficient on *EarnDec*RDoptions* (γ_6).

Sample

The sample includes observations in Compustat from 1992 to 2002 with required stock option information in ExecuComp and all data from Compustat required to calculate the control variables. The sample period begins in 1992, the first year the necessary data to estimate the stock option tax deduction is available in ExecuComp. The sample period ends in 2002 to reduce the effect of firms that began to voluntarily expense stock options when the accounting for stock options came back under scrutiny in 2002. These requirements result in a sample of 3,786 firm-year observations.

To control for the effects of outliers, we eliminate influential observations from estimating Equation (8), defined as those with studentized residuals greater than 2 (Belsley et al. 1980), and observations with ΔRD or *RDoptions* in the top 1 percent of the distribution used to estimate each regression equation. These procedures result in a sample of 3,679 observations to estimate Equation (8), of which 496 have potential earnings decreases (*EarnDec* = 1).

To estimate Equation (9), we identify credit-sensitive firm-years as those that are below the maximum θ_s and that record an income tax benefit when R&D tax credits increase. Therefore, we exclude observations for which θ_s is greater than or equal to 0.50 before considering option exercises.[15] We also exclude observations with persistent tax losses because these firms are more likely to record a deferred tax valuation allowance and less likely to record a tax benefit when R&D tax credits increase. We define firm-years with persistent tax losses as those with negative taxable income in both the current and previous year.[16] These procedures result in a sample of 2,137 credit-sensitive firm-years, of which 210 have potential earnings decreases.

---

[15] We define θ_s before considering option exercises as [(non-option R&D − Fixed Base)/non-option R&D]. Defining θ_s in this way avoids the possibility that our estimate of R&D option exercises affects our classification of firms as credit-sensitive. This definition also allows for the possibility that firms for which θ_s after option exercises is greater than 0.50, but θ_s before option exercises is less than 0.50 benefit from some or a large proportion of R&D option exercises. Results are similar if we define credit-sensitive firms using θ_s after option exercises.

IV. RESULTS

Descriptive Statistics

Table 1 describes our sample firms' R&D spending and stock option exercises over the sample period. The mean R&D spending for our sample is $190.99 million. The mean value of option exercises for non-executive employees, *Employee Exercises*, is $147.44 million and is approximately 77 percent of R&D spending. The mean R&D option exercises, *R&D Options*, is, on average, 11 percent of *Employee Exercises*. Using the *Fixed Base* column in Table 1, we estimate the effect of R&D Options on excess R&D. Before

TABLE 1
Annual Means of Inputs for the R&D Credit Calculation (in $millions)[a]

Year	Sample Size	R&D[b]	*Employee Exercises*	*R&D Options*	*Fixed Base*	*PVCR*
1992	268	129.84	125.44	2.12	124.39	0.082
1993	323	125.15	23.55	3.24	119.83	0.084
1994	299	131.49	28.40	1.94	124.18	0.099
1995	336	177.93	60.66	5.09	165.39	0.106
1996	372	151.01	107.59	9.05	150.10	0.092
1997	404	175.14	85.10	11.17	165.87	0.089
1998	384	234.26	132.21	14.65	208.67	0.085
1999	355	254.81	283.86	29.33	237.00	0.077
2000	374	236.96	417.00	50.93	236.17	0.071
2001	376	240.75	235.74	33.78	229.25	0.065
2002	295	211.08	59.54	6.06	238.15	0.058
Full Sample	3,786	190.99	147.44	16.21	184.00	0.083
Non-option R&D > Fixed Base[c]	2,136	265.09	177.56	25.52	178.93	0.143

[a] The sample includes observations in Compustat from 1992 to 2002 with required stock option information reported in ExecuComp and all data from Compustat required to calculate control variables.

[b] *R&D* equals R&D expense reported in the financial statements. *Employee Exercises* equals [(Executive Exercises/Executive Percent)*(1 − Executive Percent)] where Executive Exercises equals the value realized from options exercised by executives from ExecuComp and Executive Percent equals the ratio of the number of options granted to executives to the total number of options granted during the year. *R&D Options* equals *Employee Exercises* times the firm-specific ratio of R&D to total operating expenses averaged from 1987–2002. *Fixed Base* equals the maximum of 0.50*(R&D + *R&D Options*) and the fixed base percentage times mean sales from $t-4$ through $t-1$. The fixed base percentage equals the minimum of 16 percent and the mean ratio of R&D to sales from 1984–1988. If data are unavailable to calculate the mean ratio of R&D to sales from 1984 to 1988, then the fixed base percentage equals 3 percent. *PVCR* equals the present value of the R&D credit rate for firms with (R&D + *R&D Options*) greater than *Fixed Base*, and 0 otherwise.

[c] Sample means for firm-year observations where non-option R&D (Compustat data46) is greater than *Fixed Base*.

---

[16] We calculate current-year taxable income using the definition in Figure 1 and taxable income for the previous year using the carryback/carryforward procedure described in the Appendix. Inferences are unchanged if we define firm-years with persistent tax losses as those for which Graham's (1996) marginal tax rate is equal to 0.

considering *R&D Options*, the mean excess R&D is \$6.99 million (\$190.99 − \$184.00). When we add *R&D Options*, the mean excess R&D is \$23.20 million (\$190.99 + \$16.21 − \$184.00), which suggests that, on average, R&D Options more than triples excess R&D.

Table 2 reports descriptive statistics for the full sample (Panel A) and for credit-sensitive firm-years (Panel B). For the full sample, the mean (median) ΔRD is 0.0073 (0.0008), the mean (median) *RDoptions* is 0.0053 (0.00), and the mean (median) *Credit* is 0.0606 (0.0117). In Panel B, the mean ΔRD is 0.0016, the mean *RDoptions* is 0.0003, and the mean *Credit* is 0.0487.[17] *RD* and ΔRD are significantly lower for the credit-sensitive sample than for the remaining observations in the full sample. This difference likely results because firms with very high amounts of R&D have reached the limit on θ_s and therefore are excluded from the credit-sensitive sample. Similarly, *Credit* and *RDoptions* are significantly lower for the credit-sensitive sample. *Credit* is likely lower because firms with less R&D are less likely to have excess R&D and therefore more likely to have a credit rate equal to 0, and *RDoptions* is lower because firms with lower R&D have a lower ratio of R&D to total operating expenses used in calculating *RDoptions*.

In Table 3 we report the Pearson and Spearman correlation coefficients between the regression variables. Although we expect the correlation between the credit rate and the marginal tax rate to be positive, the correlation between *Credit* and *TaxRate* is negative (−0.047). This negative correlation results because *Credit* is defined as a negative function of *TaxRate* (*Credit* equals *PVCR**(1 − *TaxRate*)). The correlation between *PVCR* and *TaxRate* is 0.088, which is similar to the correlation of 0.06 reported by Klassen et al. (2004).[18]

Empirical Test of Hypothesis 1

In Table 4 we report OLS results from estimating Equation (8). These tests confirm prior findings using our refined measure of *Credit* and test H1. For all reported regression results, we calculate t-statistics using Huber-White (Rogers 1993) robust standard errors to correct for heteroscedasticity and serial-correlation. In Column (1) we calculate the credit rate without incorporating the carryforward feature of the credit and define *Credit* as 0.20*(1 − *TaxRate*) for observations with a tax liability greater than total R&D credits in the current or previous year, and 0 otherwise. This definition of *Credit* is similar to that reported in existing literature. Consistent with existing literature, we find that the coefficient on *Credit* is positive and significant (α_2 = 0.0010; t = 2.16), suggesting that R&D expenditures are higher for firms that receive a current benefit from the R&D credit. In Column (2) we measure *Credit* as *PVCR**(1 − *TaxRate*). Using this measure of the credit rate, the coefficient on *Credit* is also positive and significant (α_2 = 0.0188; t = 3.48), suggesting that R&D expenditures are increasing in the present value of the R&D credit rate.

In Column (3) we add an indicator variable, *EarnDec*, equal to 1 for firm-years with potential earnings decreases. Consistent with H1, the coefficient on *EarnDec* is negative and significant (α_1 = −0.0030; t = −2.83). This result suggests that firms with potential earnings decreases reduce R&D relative to other firms.

---

[17] The mean and median of *RDoptions* suggests that some skewness exists in our sample even after eliminating influential and extreme observations. This skewness results in part because some firms in our sample have zero *RDoptions*. Results are similar when we eliminate observations with zero *RDoptions*. In the "Sensitivity Tests" subsection we perform additional analysis to show that our results are not driven by a small number observations with large values for *RDoptions*.

[18] To evaluate whether our results are affected by harmful multicollinearity, we examine variance inflation factors from estimating Equation (8). All variance inflation factors are less than 3, where a variance inflation factor greater than 9 indicates harmful multicollinearity.

TABLE 2
Descriptive Statistics for Regression Variables

Panel A: Full Sample (n = 3,679)[a]

Variable[c]	Mean	Q1	Median	Q3
RD	0.0575	0.0053	0.0316	0.0876
ΔRD	0.0073	0.0000	0.0008	0.0094
Credit	0.0606	0.0000	0.0117	0.1300
RDoptions	0.0053	0.0000	0.0000	0.0008
TaxRate	0.2551	0.1750	0.3500	0.3500
Q	2.1120	1.0514	1.5377	2.4210
CFShort	−0.2748	−0.3665	−0.2322	−0.1356
IndRD	0.1801	0.0148	0.1868	0.3101
ln(GNP)	8.2307	7.5640	8.1595	9.1914
ln(Sales)	20.8941	19.7044	20.8114	22.0549
Grants	0.0479	0.0020	0.0079	0.0315
Repurchases	0.0270	0.0000	0.0013	0.0322
ROA	0.1329	0.0593	0.1125	0.1840
Assets	5,432.16	280.77	862.60	3,047.07

Panel B: Credit-Sensitive Sample[b]

	Credit-Sensitive Sample (n = 2,137)		EarnDec = 1 (n = 210)	EarnDec = 0 (n = 1,927)
Variable	**Mean**	**Median**	**Mean**	**Mean**
RD	0.0251	0.0135	0.0458	0.0228
ΔRD	0.0016	0.0000	0.0004	0.0018
Credit	0.0487	0.0000	0.0860	0.0446
RDoptions	0.0003	0.0000	0.0004	0.0003
TaxRate	0.3181	0.3500	0.2992	0.3202
Q	1.8138	1.4560	1.6894	1.8274
CFShort	−0.2216	−0.2108	−0.2460	−0.2189
IndRD	0.1321	0.0332	0.2017	0.1245
ln(GNP)	8.1857	8.1595	8.3092	8.1723
ln(Sales)	21.2698	21.1294	21.4810	21.2467
Grants	0.0145	0.0049	0.0138	0.0146
Repurchases	0.0260	0.0045	0.0318	0.0253
ROA	0.1160	0.1026	0.1241	0.1151
Assets	6,594.5600	1,105.7100	7,673.07	6,477.03

---

[a] The sample in Panel A includes observations in Compustat from 1992 to 2002 with required stock option information reported in ExecuComp and all data from Compustat required to calculate control variables. From this sample we eliminate influential observations from estimating Equation (8), defined as those with studentized residuals greater than 2, and observations with *ΔRD* or *RDoptions* in the top 1 percent of the distribution.

[b] Panel B reports descriptive statistics for the subsample of credit-sensitive firm-years and partitioned into firms with and without potential earnings decreases. Credit-sensitive firm-years are those for which excess non-option R&D (Compustat R&D minus *Fixed Base* as defined in Figure 1) is less than 0.5 of non-option R&D and for which taxable income is positive in the current or previous year. The EarnDec = 1 column reports means for credit-sensitive firm-years with earnings before extraordinary items and R&D below target earnings by an amount within year $t-1$ R&D. Target earnings equals year $t-1$ earnings before extraordinary items and R&D. The EarnDec = 0 column reports means for all other credit-sensitive firm-years.

[c] See Figure 1 for variable definitions.

TABLE 3
Pearson (Spearman) Correlations on the Upper (Lower) Diagonal
(n = 3,679)[a]

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.
1. *ΔRD*[b]		0.240	0.313	−0.152	0.296	−0.461	0.244	−0.043	−0.155	0.194	0.036	0.434
2. *Credit*	0.398		−0.028	−0.047	0.143	−0.285	0.309	−0.102	−0.052	−0.002	0.080	0.398
3. *RDoptions*	0.395	0.383		−0.256	0.456	−0.361	0.193	0.105	−0.098	0.383	0.173	0.372
4. *TaxRate*	−0.121	−0.109	−0.299		−0.040	0.100	−0.193	−0.074	0.132	−0.207	0.027	0.078
5. *Q*	0.304	0.243	0.432	0.054		−0.400	0.270	0.109	−0.074	0.268	0.229	0.544
6. *CFshort*	−0.428	−0.349	−0.437	0.074	−0.435		−0.217	−0.045	0.097	−0.208	−0.097	−0.561
7. *IndRD*	0.326	0.385	0.484	−0.249	0.271	−0.306		0.012	−0.156	0.121	0.051	0.308
8. *GNP*	−0.076	−0.101	0.161	−0.082	0.062	−0.036	0.013		0.103	0.142	0.113	−0.040
9. *ln(Sales)*	−0.149	−0.066	−0.014	0.115	−0.106	0.105	−0.206	0.094		−0.085	0.056	−0.122
10. *Grants*	0.276	0.137	0.695	−0.223	0.462	−0.362	0.244	0.333	−0.063		0.107	0.189
11. *Rep*	0.008	0.094	0.127	0.108	0.179	−0.055	0.005	0.149	0.175	0.130		0.249
12. *ROA*	0.436	0.448	0.450	0.092	0.652	−0.592	0.361	−0.070	−0.172	0.363	0.183	

---

[a] The sample includes observations in Compustat from 1992 to 2002 with required stock option information reported in ExecuComp and all data from Compustat required to calculate control variables. From this sample we eliminate influential observations from estimating Equation (8), defined as those with studentized residuals greater than 2, and observations with *ΔRD* or *RDoptions* in the top 1 percent of the distribution.

[b] See Figure 1 for variable definitions.

TABLE 4
Estimates of the Effect of Potential Earnings Decreases and the R&D Credit on R&D Spending

$$\Delta RD_{i,t} = \alpha_0 + \alpha_1 EarnDec_{i,t} + \alpha_2 Credit_{i,t} + \alpha_3 RDoptions_{i,t} + \alpha_4 RDoptions{*}Credit_{i,t} + \alpha_5 TaxRate_{i,t} + \alpha_6 Q_{i,t} + \alpha_7 CFshort_{i,t} + \alpha_8 IndRD_{i,t} + \alpha_9 GNP_t + \alpha_{10} ln(Sales)_{i,t} + \alpha_{11} Grants_{i,t} + \alpha_{12} Repurchases_{i,t} + \alpha_{13} ROA_{i,t} + \varepsilon_{i,t}.$$

Variable[a]	Predicted Sign	(1)	(2)	(3)
Intercept		0.0232***	0.0237***	0.0236***
		(4.39)	(4.29)	(4.31)
EarnDec	−			−0.0030***
				(−2.83)
Credit	+	0.0010**	0.0188***	0.0215***
		(2.16)	(3.48)	(3.95)
RDoptions	+/−	0.0796**	0.1011***	0.1065***
		(2.33)	(2.57)	(2.57)
*RDoptions*Credit*	+/−			−0.3532
				(−0.57)
TaxRate	+	−0.0181***	−0.0158***	−0.0173***
		(−5.80)	(−4.98)	(−5.23)
Q	+	0.0001	0.0001	0.0001
		(0.22)	(0.32)	(0.21)
CFshort	−	−0.0230***	−0.0228***	−0.0229***
		(−9.57)	(−9.49)	(−9.39)
IndRD	+	0.0066***	0.0062***	0.0067***
		(3.09)	(2.66)	(2.91)
GNP	+	−0.0014***	−0.0013***	−0.0013***
		(−4.07)	(−3.93)	(−3.75)
ln(Sales)	+/−	−0.0007***	−0.0007***	−0.0007***
		(−3.16)	(−3.34)	(−3.34)
Grants	+/−	0.0056	0.0061*	0.0061*
		(1.63)	(1.76)	(1.75)
Repurchases	−	−0.0215***	−0.0218***	−0.0205***
		(−3.01)	(−3.06)	(−2.93)
ROA	+	0.0396***	0.0346***	0.0347***
		(5.37)	(4.23)	(4.29)
Sample Size[b]		3,676	3,679	3,679
Adjusted R^2		0.3050	0.3015	0.3035

---

*, **, *** Significant at 0.10, 0.05, and 0.01, respectively, in two-tailed tests; t-statistics are calculated using the Huber-White procedure to correct for heteroscedasticity and serial-correlation.

[a] See Figure 1 for variable definitions. In Column (1), *Credit* equals 0.20*(1 − *TaxRate*) for firm-years with year $t-1$ or year t tax liability greater than total R&D tax credits, and 0 otherwise. In Columns (2) and (3), *Credit* equals *PVCR**(1 − *TaxRate*). *PVCR* equals the present value of the R&D credit rate. *TaxRate* equals 0.35 if the firm has (1) positive taxable income (*TI*) in year t and (2) no net operating loss (NOL) carry-forward from year $t-1$, 0.175 if the firm has (1) or (2), and 0 if the firm has both (1) and (2).

[b] The sample includes observations in Compustat from 1992 to 2002 with required stock option information reported in ExecuComp and all data from Compustat required to calculate control variables. From this sample we eliminate influential observations from estimating Equation (8), defined as those with studentized residuals greater than 2, and observations with Δ*RD* or *RDoptions* in the top 1 percent of the distribution.

The coefficient on *RDoptions* is positive and significant, which suggests that R&D is increasing in R&D option exercises. The coefficient on *RDoptions*Credit* is negative and insignificant in Column (3). This result suggests that for the full sample, the relation between the credit rate and R&D spending is not affected by R&D option exercises. We test whether this coefficient is positive and significantly higher for credit-sensitive firm-years with potential earnings decreases in Table 5.

Results related to the control variables are generally consistent with predictions. The coefficients on *CFshort*, *IndRD*, *ln(Sales)*, *Repurchases*, and *ROA* are significant and in the predicted direction. Contrary to our expectations, the coefficient on *TaxRate* is negative. Gupta et al. (2004) find a positive effect of marginal tax rates on ΔRD, however, they use the lagged value of the marginal tax rate. Our measure is similar to that in Klassen et al. (2004) and captures the current-year tax rate on incremental R&D. Consistent with our results, they find a negative and significant coefficient on *TaxRate*. They note that one possible explanation for this negative coefficient is that tax-paying firms are more mature and therefore conduct less R&D. The coefficient on Q is not significant. However, existing literature finds mixed results related to this variable. Berger (1993) finds a negative coefficient on Q in some specifications and a positive coefficient in others, Gupta et al. (2004) find a negative coefficient, and Klassen et al. (2004) report a positive coefficient on Q. The coefficient on *GNP* is negative and significant, inconsistent with our prediction, but consistent with the result in Berger (1993).[19] Consistent with existing evidence and inconsistent with a substitution effect between R&D and option grants, the coefficient on *Grants* is positive and generally significant.

Empirical Test of Hypothesis 2

In Table 5 we report the results from estimating Equation (9) for the subsample of credit-sensitive firm-years. In Column (1) we estimate Equation (9) excluding the interactions with *EarnDec*. Consistent with the results in Table 4, the coefficient on *EarnDec* is negative and significant ($\gamma_1 = -0.0026$; $t = -3.99$), and the coefficient on *Credit* is positive and significant ($\gamma_2 = 0.0198$; $t = 6.53$). Although we expect the coefficient on *Credit* to be increasing in *RDoptions* only for firms with financial constraints, the coefficient on *RDoptions*Credit* is positive and significant for the full sample of credit-sensitive firm-years ($\gamma_4 = 8.9404$; $t = 3.13$). This result suggests that a sufficient number of observations in the credit-sensitive sample face financial constraints, and that the positive effect of the credit rate on R&D is increasing in R&D option exercises.[20]

In Column (2) of Table 5 we report the results of estimating Equation (9) with the *EarnDec* interactions. Consistent with the results in Column (1), the coefficient on *EarnDec* is negative and significant ($\gamma_1 = -0.0027$; $t = -2.69$), the coefficient on *Credit* is positive and significant ($\gamma_2 = 0.0184$; $t = 5.89$), and the coefficient on *RDoptions*Credit* is positive and significant ($\gamma_4 = 6.5721$; $t = 2.25$). The coefficient on *EarnDec*Credit* is not significant,

---

[19] We also estimate Equation (8) using the change in GNP from year $t-1$ to t and find a positive but insignificant coefficient on this variable. The coefficient is also negative and significant if we include the natural log of GDP or GNP in year $t-1$. All other inferences are unchanged.

[20] The results for *RDoptions* and *RDoptions*Credit* differ from those in Table 4. One possible explanation is that the full sample includes firms with very high R&D, as evidenced by the fact that they have reached the maximum θ of 0.50 (while firms that have maxed out θ are excluded from the credit-sensitive sample), and also use options extensively to compensate employees, thus causing a positive coefficient on *RDoptions*. The coefficient on *RDoptions*Credit* is likely significant in Table 5 and not in Table 4 because the credit-sensitive subsample is designed to maximize the sensitivity of the sample to this variable, i.e., in forming the credit-sensitive sample we eliminate observations that we expect to be insensitive to R&D credits from R&D option exercises.

TABLE 5
Estimates of the Effect of R&D Option Exercises and Potential Earnings Decreases on R&D Spending

$$\begin{aligned}\Delta RD_{i,t} = {} & \gamma_0 + \gamma_1 EarnDec_{i,t} + \gamma_2 Credit_{i,t} + \gamma_3 RDoptions_{i,t} + \gamma_4 RDoptions{*}Credit_{i,t} \\ & + \gamma_5 EarnDec{*}Credit_{i,t} + \gamma_6 EarnDec{*}RDoptions_{i,t} \\ & + \gamma_7 EarnDec{*}RDoptions{*}Credit_{i,t} + \gamma_8 TaxRate_{i,t} + \gamma_9 Q_{i,t} \\ & + \gamma_{10} CFshort_{i,t} + \gamma_{11} IndRD_{i,t} + \gamma_{12} GNP_t + \gamma_{13} ln(Sales)_{i,t} \\ & + \gamma_{14} Grants_{i,t} + \gamma_{15} Repurchases_{i,t} + \gamma_{16} ROA_{i,t} + \eta_{i,t}.\end{aligned}$$

		(1)		(2)	
Variable[a]	**Pred. Sign**	**Coefficient**	**t-statistic**	**Coefficient**	**t-statistic**
Intercept		0.0069**	2.19	0.0063**	1.97
EarnDec	−	−0.0026***	−3.99	−0.0027***	−2.69
Credit	+	0.0198***	6.53	0.0184***	5.89
RDoptions	+/−	0.3419	1.02	0.5354	1.53
*RDoptions*Credit*	+/−	8.9404***	3.13	6.5721**	2.25
*EarnDec*Credit*	+/−			0.0046	0.48
*EarnDec*RDoptions*	−			−6.4201**	−2.45
*EarnDec*RDoptions*Credit*	+			49.2464***	2.69
TaxRate	+	−0.0046**	−2.15	−0.0040*	−1.89
Q	+	−0.0004***	−2.59	−0.0004***	−2.60
CFshort	−	−0.0055***	−5.14	−0.0056***	−5.21
IndRD	+	−0.0014	−1.00	−0.0011	−0.82
GNP	+	−0.0007***	−4.23	−0.0007***	−4.31
ln(Sales)	+/−	0.0000	−0.32	0.0000	−0.13
Grants	+/−	0.0017	0.69	0.0012	0.50
Repurchases	−	−0.0024	−0.91	−0.0020	−0.78
ROA	+	0.0145***	3.39	0.0149***	3.40
Sample Size[b]		2,137		2,137	
Adjusted R^2		0.2010		0.2070	

---

*, **, *** Significant at 0.10, 0.05, and 0.01, respectively, in two-tailed tests; t-statistics are calculated using the Huber-White procedure to correct for heteroscedasticity and serial-correlation.

[a] See Figure 1 for variable definitions.

[b] The sample includes credit-sensitive firm-years in Compustat from 1992 to 2002 with required stock option information reported in ExecuComp and all data from Compustat required to calculate control variables. Credit-sensitive firm-years are those for which excess non-option R&D (Compustat R&D minus *Fixed Base* as defined in Figure 1) is less than 0.5 of non-option R&D and taxable income is positive in the current or previous year. From this sample we eliminate influential observations from estimating Equation (8), defined as those with studentized residuals greater than 2, and observations with ΔRD or *RDoptions* in the top 1 percent of the distribution.

which likely reflects the ambiguous nature of our predictions for this variable. The coefficient on *EarnDec*RDoptions* is negative and significant consistent with our prediction that firms with potential earnings decreases are more likely to decrease R&D to offset the dilutive effect of option exercises.

Hypothesis 2 predicts that R&D option exercises increase the positive relation between the credit rate and R&D significantly more for firms with potential earnings decreases than for other firms. In support of this hypothesis, the coefficient on *EarnDec*RDoptions*Credit* is positive and significant ($\gamma_7 = 49.2464$; $t = 2.69$). This result provides support for H2 and suggests that, among credit-sensitive firms, the effect of short-term earnings constraints on changes in R&D is decreasing in tax credits generated by R&D option exercises. Further, this finding is consistent with our expectation that tax credits generated by R&D option exercises mitigate myopic R&D investment.[21,22]

Sensitivity Tests

Alternate Measures of R&D Option Exercises

Our main tests use *R&D Options*, the tax deduction from employee option exercises (*Employee Exercises*) multiplied by the ratio of R&D to operating expenses, to estimate R&D option exercises. This estimate assumes that, across firms, the ratio of R&D option exercises to total non-executive option exercises varies proportionately with the ratio of R&D to operating expense. This estimate of *R&D Options* results in very small numbers because total operating expenses are large relative to R&D expense. It also likely makes *R&D Options* a conservative estimate of R&D option exercises because firms exercise discretion in determining which expenses are R&D-related for tax purposes (Hall and Van Reenen 2000) and because existing studies find that the use of stock options to compensate non-executive employees is increasing in R&D (Core and Guay 2001). If *RDoptions*, on average, underestimates the underlying construct, then this downward bias will increase the magnitude of our coefficient estimates.[23] To determine the reasonableness of our results in Table 5, and whether these results are driven by our assumptions, we re-estimate Equation (9) using two additional estimates of this construct.

First, we use the tax deduction from total non-executive options divided by assets to estimate *RDoptions*, and we report the results in Column (1) of Table 6. This measure overestimates R&D option exercises, but represents a firm's ability to include the tax deduction from stock option exercises in qualified R&D expenditures. It also results in much larger estimates of *RDoptions*, with a mean of 0.54 percent of assets and 21.5 percent of non-option R&D. Consistent with the results in Table 5, the coefficient on *EarnDec*RDoptions* is negative, but it is insignificant ($\gamma_6 = -0.2565$; $t = -1.46$). Consistent with H2, the coefficient on *EarnDec*RDoptions*Credit* is positive and significant ($\gamma_7 = 2.6227$; $t = 1.76$).

---

[21] Results are similar when we replicate the analyses in Table 5 using Huber-White standard errors clustered by firm and include year dummies. This procedure corrects for heteroscedasticity, serial-correlation, and cross-sectional correlation. Inferences are unchanged when we include influential observations and when we use the Newey and West (1987) procedure to correct for heteroscedasticity and serial-correlation. Inferences are also similar when we include the fixed base amount divided by beginning of year total assets as an additional control variable, and when we include income before extraordinary items, R&D, and depreciation divided by beginning-of-year total assets, the measure of funds available to finance R&D used in Berger (1993) and Gupta et al. (2004), in place of *CFshort*. Results are also not sensitive to using Graham's (1996) marginal tax rates in place of the trichotomous tax rate measure or to including the percentage of firms held by institutional investors as an additional control variable.

[22] We evaluate the effect of forecasting taxable income by re-estimating Equation (9) using the estimate of *Credit* in Column (1) of Table 4, set equal to $0.2*(1 - TaxRate)$ if the firm has a tax liability greater than the R&D credit in the current or previous year, and 0 otherwise. We find that the coefficient on *Credit* is 0.0194 ($t = 5.89$), the coefficient on *RDoptions*Credit* is 4.2337 ($t = 1.47$), and the coefficient on *EarnDec*RDoptions*Credit* is 4.0484 ($t = 0.35$). These results suggest that forecasting the taxable income stream improves the power of our empirical tests.

[23] Determining the precise amount of R&D option exercises is difficult since firms rarely report this amount on their financial statements during our sample period. As an example of the potential magnitude of R&D option exercises, consider the *Apple Computer vs. The Commissioner* Tax Court decision that formally established the tax deduction from R&D-related stock option exercises as a qualified R&D expenditure. In this case, the disputed amount was $18.454 million in R&D-related stock option deductions for the 1983 tax year, and Apple reported $60 million in R&D expense on its financial statements for the same year.

TABLE 6
Estimates of the Effect of R&D Option Exercises and Potential Earnings Decreases on R&D Spending Using Alternate Definitions of R&D Option Exercises

$$\begin{aligned}\Delta RD_{i,t} = {} & \gamma_0 + \gamma_1 EarnDec_{i,t} + \gamma_2 Credit_{i,t} + \gamma_3 RDoptions_{i,t} + \gamma_4 RDoptions{*}Credit_{i,t} \\ & + \gamma_5 EarnDec{*}Credit_{i,t} + \gamma_6 EarnDec{*}RDoptions_{i,t} \\ & + \gamma_7 EarnDec{*}RDoptions{*}Credit_{i,t} + \gamma_8 TaxRate_{i,t} + \gamma_9 Q_{i,t} \\ & + \gamma_{10} CFshort_{i,t} + \gamma_{11} IndRD_{i,t} + \gamma_{12} GNP_t + \gamma_{13} ln(Sales)_{i,t} \\ & + \gamma_{14} Grants_{i,t} + \gamma_{15} Repurchases_{i,t} + \gamma_{16} ROA_{i,t} + \eta_{i,t}.\end{aligned}$$

		(1) Non-Executive Options		(2) Ratio of R&D to SGA	
Variable[a]	**Pred. Sign**	**Coefficient**	**t-statistic**	**Coefficient**	**t-statistic**
Intercept		0.0061*	1.88	0.0068**	2.09
EarnDec	−	−0.0030***	−2.80	−0.0031***	−2.88
Credit	+	0.0190***	5.58	0.0185***	5.86
RDoptions	+/−	0.0369**	2.38	0.3018***	3.66
*RDoptions*Credit*	+/−	0.0665	0.35	0.3932	0.45
*EarnDec*Credit*	+/−	0.0039	0.36	0.0046	0.44
*EarnDec*RDoptions*	−	−0.2565	−1.46	−1.1157	−1.41
*EarnDec*RDoptions*Credit*	+	2.6227*	1.76	12.0192**	2.25
TaxRate	+	−0.0048**	−2.15	−0.0044**	−2.03
Q	+	−0.0004**	−2.10	−0.0003**	−2.04
CFshort	−	−0.0057***	−4.92	−0.0055***	−5.11
IndRD	+	−0.0007	−0.49	−0.0007	−0.52
GNP	+	−0.0008***	−4.22	−0.0008***	−4.61
ln(Sales)	+/−	0.0000	−0.03	0.0000	−0.11
Grants	+/−	0.0009	0.42	0.0006	0.31
Repurchases	−	−0.0030	−1.07	−0.0020	−0.77
ROA	+	0.0174***	3.72	0.0153***	3.38
Sample Size[b]		2,134		2,137	
Adjusted R^2		0.1878		0.1978	

*, **, *** Significant at 0.10, 0.05, and 0.01, respectively, in two-tailed tests; t-statistics are calculated using the Huber-White procedure to correct for heteroscedasticity and serial-correlation.

[a] In Column (1) *RDoptions* equals *Employee Exercises* divided by assets at the end of year $t-1$. *Employee Exercises* equals [(Executive Exercises/Executive Percent)*(1 − Executive Percent)], Executive Exercises equals the value realized from options exercised by executives, and Executive Percent equals the ratio of the number of options granted to executives to the total number of options granted during the year. In Column (2) *RDoptions* equals *Employee Exercises* times the mean ratio of R&D to selling, general, and administrative expenses from 1987 to 2002, divided by total assets at the end of year $t-1$. All other variables are defined in Figure 1.

[b] The sample includes credit-sensitive firm-years in Compustat from 1992 to 2002 with required stock option information reported in ExecuComp and all data from Compustat required to calculate control variables. Credit-sensitive firm-years are those for which excess non-option R&D (Compustat R&D minus *Fixed Base* as defined in Figure 1) is less than 0.5 of non-option R&D and taxable income is positive in the current or previous year. From this sample we eliminate influential observations from estimating Equation (8), defined as those with studentized residuals greater than 2, and observations with ΔRD or *RDoptions* in the top 1 percent of the distribution.

Second, we estimate *RDoptions* as *Employee Exercises* times the firm-specific ratio of R&D to selling, general, and administrative expenses averaged from 1987 through 2002, and we report the results in Column (2) of Table 6.[24] Consistent with the results in Column (1), the coefficient on *EarnDec*RDoptions* is negative and insignificant ($\gamma_6 = -1.1157$; $t = -1.41$), and the coefficient on *EarnDec*RDoptions*Credit* is positive and significant ($\gamma_7 = 12.0192$; $t = 2.25$). These results provide support for the reliability of this measure and additional support for our hypotheses.

To obtain a large, generalizable sample, we use data from ExecuComp to estimate the total tax deduction from stock option exercises. Our estimate of R&D option exercises assumes that the percentage of total options granted to executives equals the percentage of total options exercised by executives. We test whether our results are sensitive to this assumption by replicating the tests in Table 5 for a reduced sample for which we calculate the tax deduction from stock option exercises using detailed information from the financial statement footnotes. This sample includes firms with positive values of R&D reported in Compustat for each year from 1984 to 1988 and 1997 to 2002, required stock option information in the financial statement footnotes, and all required data to calculate control variables. Following Graham et al. (2004), we calculate the total tax deduction from stock option exercises as the number of options exercised during the year, times the spread between the weighted average strike price for options granted during the year (an estimate of the fair market value at exercise) and the weighted average strike price for options exercised during the year (an estimate of the exercise price).[25] We multiply this measure by the firm-specific ratio of R&D to total operating expenses, averaged from 1987 through 2002, and divide by beginning of year total assets to get *RDoptions*.

Table 7 reports results from estimating Equation (9) for 416 firm-years in both the footnote sample and the credit-sensitive sample. In the ExecuComp (Footnote) column, we report results using the ExecuComp (Footnote) measure of *RDoptions*. Inconsistent with the results in Table 5, the coefficient on *EarnDec* is not significantly different from zero in either the ExecuComp or Footnote columns. However, when we estimate Equation (9) without the *EarnDec* interactions, the coefficient on *EarnDec* is negative and significant ($\gamma_1 = -0.0032$, $t = -2.50$ using the ExecuComp calculation of *RDoptions* and $\gamma_1 = -0.0033$, $t = -2.52$ using the footnote calculation of *RDoptions*) consistent with the results in Column (1) of Table 5. Thus, we find evidence using this reduced sample that firms decrease R&D to avoid earnings decreases.

The coefficient on *EarnDec*RDoptions* is negative and significant in both columns ($\gamma_6 = -7.4375$, $t = -1.95$ in the ExecuComp column; $\gamma_6 = -4.7384$, $t = -2.75$ in the Footnote column) similar to Table 5. Also consistent with the results in Table 5, the coefficient on *Credit* is positive and significant in both columns ($\gamma_2 = 0.0143$, $t = 2.46$ in the ExecuComp column; $\gamma_2 = 0.0174$, $t = 3.03$ in the Footnote column) and the coefficient on *EarnDec*RDoptions*Credit* is positive and significant in both columns ($\gamma_7 = 49.3848$, $t = 1.77$ in the ExecuComp column; $\gamma_7 = 32.2326$, $t = 2.58$ in the Footnote column). These results provide evidence that our findings are not sensitive to using ExecuComp data to estimate *RDoptions*.

---

[24] If the average ratio of R&D to selling, general, and administrative expenses from 1987 through 2002 is greater than 1, then we use the industry mean ratio of R&D to selling, general, and administrative expenses over the same period. The mean *RDoptions* for this definition is 0.0009.

[25] When this calculation produces a negative number we replace the weighted average exercise price of options granted with the fiscal year price high.

TABLE 7
Estimates of the Effect of R&D Option Exercises and Potential Earnings Decreases on R&D Spending Using Detailed Footnote Data

$$\begin{aligned}\Delta RD_{i,t} = {} & \gamma_0 + \gamma_1 EarnDec_{i,t} + \gamma_2 Credit_{i,t} + \gamma_3 RDoptions_{i,t} + \gamma_4 RDoptions{*}Credit_{i,t} \\ & + \gamma_5 EarnDec{*}Credit_{i,t} + \gamma_6 EarnDec{*}RDoptions_{i,t} \\ & + \gamma_7 EarnDec{*}RDoptions{*}Credit_{i,t} + \gamma_8 TaxRate_{i,t} + \gamma_9 Q_{i,t} \\ & + \gamma_{10} CFshort_{i,t} + \gamma_{11} IndRD_{i,t} + \gamma_{12} GNP_t + \gamma_{13} ln(Sales)_{i,t} \\ & + \gamma_{14} Grants_{i,t} + \gamma_{15} Repurchases_{i,t} + \gamma_{16} ROA_{i,t} + \eta_{i,t}.\end{aligned}$$

		(1) ExecuComp Data		(2) Footnote Data	
Variable[a]	Pred. Sign	Coefficient	t-statistic	Coefficient	t-statistic
Intercept		0.0007	0.08	−0.0004	−0.04
EarnDec	−	−0.0018	−1.15	−0.0022	−1.49
Credit	+	0.0143**	2.46	0.0174***	3.03
RDoptions	+/−	0.0031	0.01	0.1145	0.21
*RDoptions*Credit*	+/−	3.4339	0.99	0.0744	0.02
*EarnDec*Credit*	+/−	−0.0019	−0.12	−0.0012	−0.07
*EarnDec*RDoptions*	−	−7.4375*	−1.95	−4.7384***	−2.75
*EarnDec*RDoptions*Credit*	+	49.3848*	1.77	32.2326**	2.58
TaxRate	+	0.0026	0.51	−0.0013	−0.22
Q	+	−0.0002	−0.38	0.0000	−0.02
CFshort	−	−0.0137***	−3.32	−0.0142***	−3.31
IndRD	+	−0.0006	−0.31	−0.0007	−0.32
GNP	+	−0.0018**	−2.14	−0.0016*	−1.87
ln(Sales)	+/−	0.0005*	1.75	0.0005*	1.75
Grants	+/−	0.0016	0.62	0.0021	0.76
Repurchases	−	0.0087	1.12	0.0080	0.96
ROA	+	0.0154	1.14	0.0194	1.29
Sample Size[b]		416		416	
Adjusted R^2		0.2419		0.2241	

*, **, *** Significant at 0.10, 0.05, and 0.01, respectively, in two-tailed tests; t-statistics are calculated using the Huber-White procedure to correct for heteroscedasticity and serial-correlation.

[a] See Figure 1 for variable definitions.

[b] The sample includes credit-sensitive firm-years from 1997–2002 with positive R&D for each year 1984–1988 and 1997–2002, and all required data from financial statement footnotes, Compustat, and ExecuComp. Credit-sensitive firm-years are those for which excess non-option R&D (Compustat R&D minus *Fixed Base* as defined in Figure 1) is less than 0.5 of non-option R&D and taxable income is positive in the current or previous year.

Skewness in R&D Option Exercises

To reduce the possibility that our results are driven by a few large observations, we eliminate influential observations and observations in the top 1 percent of ΔRD and *RDoptions* for our main tests. However, in Panel B of Table 2, the mean, median, and 75th percentile of *RDoptions* indicate that this measure remains highly skewed even after eliminating influential and extreme observations. To determine whether our results are driven by a few large observations, we eliminate observations with *RDoptions* greater than the

90th percentile. Inferences are unchanged. We also replace *RDoptions* with a dummy variable equal to 1 for observations with *RDoptions* greater than the 90th percentile. The coefficient on *EarnDec*RDoptions*Credit* becomes insignificant (γ_7 = 0.0358; t = 0.79), suggesting that the results are not driven by these firms. When we replace *RDoptions* with a dummy variable equal to 1 for observations with *RDoptions* greater than the 75th percentile, the coefficient on *EarnDec*RDoptions*Credit* is again positive and significant (γ_7 = 0.0469; t = 1.87), suggesting that our results are not driven by a few large observations.

Other Sensitivity Tests

In additional sensitivity tests, we control for factors unrelated to R&D tax credits that may be correlated with both stock option exercises and R&D spending. First, a correlation between industry factors and option exercises could contribute to the results we observe. To control for this possibility, we include the industry mean R&D as a control variable in Equations (8) and (9). As an additional sensitivity test, we replace *IndRD* with dummy variables that correspond to Fama and French (1997) industry classifications and find similar results. To control for the possibility that firms both experience higher option exercises and increase R&D in good years, we include *ROA* as a control variable in Equations (8) and (9). To further control for this possibility, we include the firm's stock return from the beginning of year $t-2$ through the end of year t as an additional control variable and again find similar results.

To mitigate the possibility that the results we observe reflect a correlation between error in measuring management's expectation of employee option exercises and variables of interest, we include measures of stock returns and unexpected stock returns as additional control variables. To the extent that unexpected exercises are related to unexpected changes in stock prices, these controls reduce the effect of any correlation between measurement error in anticipated employee option exercises and our variables of interest. In separate tests, we include the year t stock return, the difference between year t and year $t-1$ stock return, the stock return in the fourth quarter of year t, and the difference between the stock return in the fourth quarter of year t and the average stock return in the first three quarters of year t. Inferences are unchanged for each of these measures, providing additional support for our hypotheses.

Alternative Explanation

One potential alternative explanation for the results we observe is that *Credit* is a proxy for the importance of R&D to the firm rather than the tax benefits of R&D spending. Based on this view, firms for which it is more economically costly to reduce R&D, i.e., firms with higher amounts of R&D and therefore higher values for *Credit*, will be less likely to reduce R&D to meet short-term earnings constraints than other firms.[26] However, our empirical design helps control for this possibility in several ways. First, we include Tobin's q to control for the firm's investment opportunity set because we expect firms with more investment opportunities to invest in more R&D. Thus, to the extent that firms with more investment opportunities face a greater economic cost of reducing R&D, this variable helps alleviate the concern that *Credit* is a proxy for the importance of R&D. Second, as a sensitivity test we estimate the empirical model with R&D as the dependent variable and year $t-1$ R&D as an independent variable and find similar results. Lagged R&D helps to control for the stickiness of R&D and therefore the economic cost of reducing R&D.

---

[26] The correlation between *RD* and *Credit* is 0.379.

Finally, if *Credit* measures the economic cost of reducing R&D and therefore firms with higher *Credit* are less likely to reduce R&D to meet an earnings target, then we would expect the coefficient on *EarnDec*Credit* to be positive. However, in Table 5 this coefficient is not significantly different from zero. Further, we would not expect the magnitude of this effect to vary with R&D option exercises. Our main test shows that R&D option exercises decrease the effect of short-term earnings constraints on the sensitivity of R&D to the credit rate. Because R&D option exercises measure the earnings effect of R&D tax credits, the interaction with R&D option exercises more closely ties *Credit* to tax-related effects and further alleviates the concern that our results reflect that *Credit* is a proxy for the importance of R&D rather than the tax effect of R&D.

Economic Significance

To gauge the importance of tax credits from R&D option exercises in mitigating myopic R&D investment, we estimate the extent to which the positive effect of tax credits from R&D option exercises offsets the negative effect of potential earnings decreases on R&D spending. The coefficient of −0.0027 on *EarnDec* suggests that credit-sensitive firm-years with zero R&D option exercises, a zero R&D tax credit rate, and a potential earnings decrease reduce R&D relative to similar firms without a potential earnings decrease by 0.27 percent of assets. As evidenced by the coefficient of −6.4201 on *EarnDec*RDoptions*, R&D option exercises increase the magnitude of this effect. This result is consistent with our prediction and with Bens et al. (2002) who find that firms reduce R&D to offset the dilutive effect of option exercises and also suggest that option exercises play an important role in myopic R&D investment. Assuming *RDoptions* equals the sample mean of 0.0003 and *Credit* equals zero, firms with a potential earnings decrease reduce R&D relative to other firms by 0.46 percent of assets [−0.0027 + (−6.4201*0.0003)]. The coefficient of 49.2464 on *EarnDec*RDoptions*Credit* suggests that tax credits from R&D option exercises reduce the magnitude of the negative effect of potential earnings decreases on R&D spending. Assuming *RDoptions* and *Credit* equal the sample means of 0.0003 and 0.0487, respectively, tax credits from R&D option exercises increase R&D spending for firms with potential earnings decreases by 0.07 percent of assets (49.2464*0.0003*0.0487 = 0.0007), thereby offsetting the 0.46 percent decrease by 16 percent (0.0007/0.0046). If we assume the firm is fully taxable (i.e., *Credit* equals 0.13),[27] tax credits from R&D options increase R&D spending by 0.18 percent of assets (49.2464*0.0003*0.13 = 0.0018), thereby offsetting the 0.46 percent decrease by 42 percent (0.0018/0.0046). Thus, the results of this study suggest that tax credits from R&D option exercises play an important role in mitigating myopic R&D investment.[28]

V. SUMMARY AND CONCLUSION

The research on myopic R&D investment finds that firms decrease R&D spending to meet earnings benchmarks. However, this research does not consider R&D tax credits from stock option exercises, which reduce the financial reporting costs of R&D spending. Specifically, when R&D-related employees exercise stock options, the tax deduction from these option exercises increase R&D tax credits generated by non-option R&D. These tax credits increase earnings while the exercises themselves do not directly affect earnings, decreasing the after-tax earnings effect of R&D. Because of this decrease in the financial reporting costs of R&D spending, we expect that, as tax credits from R&D-related stock option exercises increase, firms near an earnings benchmark can invest in more R&D before reaching the constraint. As a result, they are less likely to forgo or postpone R&D spending to avoid missing the earnings target. Therefore, we investigate whether R&D tax credits generated by stock option exercises mitigate myopic R&D investment.

---

[27] For a fully taxable firm with an R&D credit rate of 20 percent and marginal tax rate of 35 percent, *Credit* = [0.20*(1 − 0.35)] = 0.13.

[28] The estimates of economic significance are of similar magnitudes when we use the two alternate definitions of *RDoptions*. For example, when we define *RDoptions* as total non-executive option exercises, the mean *RDoptions* is 0.0054. The coefficient on *EarnDec* in Table 6 Column (1) is −0.0030, the coefficient on *EarnDec*RDoptions* is −0.2565, and the coefficient on *EarnDec*RDoptions*Credit* is 2.6227. Therefore, firms with a potential earnings decrease, the mean *RDoptions* of 0.0054, and zero *Credit* decrease R&D relative to similar firms without a potential earnings decrease by 0.44 percent of assets [−0.0030 + (−0.2656*0.0054)]. For a fully taxable firm, tax credits from R&D option exercises offset this decrease by 0.18 percent of assets (2.6227*0.0054*0.13) or 41 percent (0.0018/0.0044) for a fully taxable firm.

We find that firms with potential earnings decreases, which we define as firms for which current-year earnings before R&D is less than earnings before R&D in the previous year, reduce R&D expenditures relative to other firms and that R&D spending is positively related to the R&D credit rate, consistent with existing research on myopic R&D investment and the effectiveness of R&D tax incentives. Consistent with predictions from our theoretical model, we also find that the positive relation between the credit rate and R&D is increasing in R&D-related stock option exercises, and the magnitude of this effect is greater for firms with potential earnings decreases. Our results suggest that R&D tax credits generated by stock option exercises mitigate myopic R&D investment. Specifically, we find that firms with R&D option exercises that face a potential earnings decrease reduce R&D relative to other firms by 0.46 percent of total assets, but that R&D tax credits generated by stock option exercises offset this decrease by 16 percent on average and by as much as 42 percent for a fully taxable firm. The results of this study provide useful information that can help researchers and investors identify firms that are more or less likely to engage in myopic R&D investment and suggest that R&D tax credits from stock option exercises can be an important component of the R&D tax credit for firms with large amounts of R&D-related stock option exercises.

APPENDIX

ESTIMATE OF THE CREDIT RATE ON INCREMENTAL R&D SPENDING

To calculate the present value of the R&D credit rate we (1) generate a stream of taxable income (*TI*) consisting of an estimate of *TI* from the financial statements for periods −2 to 0 and forecasted *TI* for periods 1 to 20, (2) carry any resulting net operating loss back two years and forward 20 years, (3) determine the period in which the firm has a tax liability greater than the R&D credit, and (4) discount the statutory rate of the credit from the period determined in step (3) to period 0.

We calculate estimated and forecasted taxable income for periods −2 to 20 (*TI*) by employing the method developed by Shevlin (1990), Manzon (1994), Graham (1996), and Graham et al. (2004). We estimate *TI* as:

$$TI_{i,t} = PTI_{i,t} - (DTxp_{i,t}/tr)_{i,t} - TotalOptions_{i,t} \tag{A1}$$

where *PTI* is reported book pre-tax income, *DTxp* is the deferred tax expense from the statement of cash flows, *tr* is the statutory tax rate of 0.35, and *TotalOptions* equals the tax deduction from stock option exercises by executives divided by the ratio of options granted to executives to total options granted during the year.[29] Forecasts of *TI* assume that taxable income follows a random walk with drift (Shevlin 1990; Graham 1996):

$$TI_{i,t} = TI_{i,t-1} + \delta_i + \varepsilon_{i,t} \tag{A2}$$

where δ_i is a drift term equal to $\overline{\Delta TI_i}$ and $\varepsilon_{i,t}$ is an error term drawn as a random number from a distribution with mean zero and variance $\sigma_i^2(\Delta TI)$. We obtain the parameters δ_i and $\sigma_i^2(\Delta TI)$ using estimates of the first difference in *TI* from Equation (A1) over the prior three years.[30] We generate an income stream from period -2 to 20 by setting *TI* for periods -2 to 0 equal to the estimate from Equation (A1) and using Equation (A2) to forecast *TI* for periods 1 to 20.

We adjust the income stream to account for the ability to carry net operating losses back one year and forward 20 years.[31] From the adjusted income stream, we estimate the period in which the tax liability on *TI* is greater than the R&D credit, where the R&D credit equals 0.20*(*Non-Option R&D* + *R&D Options* − *FixedBase*), *Non-Option R&D* equals R&D reported in Compustat, and *FixedBase* equals the maximum of 50 percent of current year non-option R&D plus *R&D Options* and the firm's fixed base percentage times mean sales from $t-4$ through $t-1$. For firms with R&D during the 1984–1988 fixed base period, the fixed base percentage equals the minimum of 16 percent and the mean ratio of R&D to sales over the fixed base period; otherwise, the fixed base percentage equals 3 percent. We discount the statutory rate of the credit back from this period using a 7 percent discount rate.[32]

If the adjusted income stream indicates that the firm will not have a tax liability greater than the credit during the carryback or carryforward periods, then the discounted credit rate equals zero. We repeat the simulation procedure 50 times for each firm-year, each using different random number draws from a normal distribution with mean zero and variance $\sigma_i^2(\Delta TI)$ to forecast a new income stream. We compute the average of the 50 estimates of the discounted rate of credit to obtain one firm-year observation of the present value of the credit rate (*PVCR*). Because we define *TI* after deducting *R&D* and *Total Options*, *PVCR* captures the tax benefit from the R&D credit on the last dollar of R&D spending.

REFERENCES

Altshuler, R. 1988. A dynamic analysis of the research and experimentation tax credit. *National Tax Journal* 41: 453–466.

---

[29] If deferred tax expense from the statement of cash flows is not reported in Compustat, we estimate the deferred tax expense as the change in ending deferred taxes from the balance sheet. In this case we calculate *TI* as $TI_{i,t} = PTI_{i,t} -/- (\Delta DTxp_{i,t}/tr)_{i,t} - TotalOptions_{i,t}$, where $\Delta Dtxp$ is the annual change in deferred taxes reported on the balance sheet.

[30] Due to data constraints, we use *TI* before option exercises to calculate the drift term and assume that *TotalOptions* and *TI* before stock options follow similar random walk patterns.

[31] The net operating loss (NOL) and R&D credit carryforward periods changed in 1997. During our sample period, from 1992 through 1997, R&D credits and NOLs could be carried back three years and forward 15 years. From 1998 through 2002, R&D credits could be carried back one year and forward 20 years, and NOLs could be carried back two years and forward 20 years. Our calculations incorporate these changes. The R&D credit rate remains constant at 0.20 during the sample period.

[32] The 7 percent discount rate is the average Moody's corporate bond yield during the sample period and is the same source used by Graham (1996). As a sensitivity test we use the annual mean of the implied cost of equity capital calculated using the method in Gebhardt et al. (2001) as reported in Dhaliwal et al. (2005) and Dhaliwal et al. (2007) to discount the credit rate and find similar results. Inferences are also unchanged if we do not discount the credit rate.

Baber, W., P. Fairfield, and J. Haggard. 1991. The effect of concern about reported income on discretionary spending decisions: The case of research and development. *The Accounting Review* 66 (4): 818–829.

Belsley, D. A., E. Kuh, and R. E. Welsch. 1980. *Regression Diagnostics*. New York, NY: John Wiley & Sons.

Bens, D. A., V. Nagar, and M. H. F. Wong. 2002. Real investment implications of employee stock option exercises. *Journal of Accounting Research* 40 (2): 359–393.

Berger, P. G. 1993. Explicit and implicit tax effects of the R&D tax credit. *Journal of Accounting Research* 31: 131–172.

Bushee, B. J. 1998. The influence of institutional investors on myopic R&D investment behavior. *The Accounting Review* 73 (3): 305–333.

Cheng, S. 2004. R&D expenditures and CEO compensation. *The Accounting Review* 79 (2): 305–328.

Cohen, D., A. Dey, and T. Lys. 2008. Real and accrual-based earnings management in the pre- and post-Sarbanes-Oxley periods. *The Accounting Review* 83 (3): 757–787.

Core, J., and W. Guay. 2001. Stock option plans for non-executive employees. *Journal of Financial Economics* 61: 253–287.

Dhaliwal, D., C. Gleason, and L. Mills. 2004. Last chance earnings management: Using the tax expense to meet analysts' forecasts. *Contemporary Accounting Research* 21 (2): 431–459.

———, L. Krull, O. Z. Li, and W. Moser. 2005. Dividend taxes and implied cost of equity capital. *Journal of Accounting Research* 43 (5): 675–708.

———, ———, and ———. 2007. Did the 2003 tax act reduce the cost of equity capital? *Journal of Accounting and Economics* 43 (1): 121–150.

Eisner, R., S. Albert, and M. Sullivan. 1984. The new incremental tax credit for R&D: Incentive or disincentive? *National Tax Journal* 37: 171–183.

Fama, E., and K. French. 1997. Industry costs of equity. *Journal of Financial Economics* 43: 153–193.

Fields, T. D., T. Z. Lys, and L. Vincent. 2001. Empirical research on accounting choice. *Journal of Accounting and Economics* 31: 255–307.

Gebhardt, L., C. Lee, and B. Swaminathan. 2001. Toward an implied cost of capital. *Journal of Accounting Research* 39: 135–176.

Graham, J. 1996. Debt and the marginal tax rate. *Journal of Financial Economics* 41: 41–74.

———, M. Lang, and D. Shackelford. 2004. Employee stock options, corporate taxes and debt policy. *Journal of Finance* 59 (1): 585–1618.

Gunny, K. 2005. What are the consequences of real earnings management? Working paper, University of Colorado.

Gupta, S., Y. Hwang, and A. Schmidt. 2004. R&D spending fools? An analysis of the R&D credit's incentive effects after the Omnibus Budget Reconciliation Act of 1989. Working paper, Arizona State University and Columbia University.

Hall, B. 1993. R&D tax policy during the eighties: Success or failure? *Tax Policy and the Economy* 7: 1–36.

———, and J. Van Reenen. 2000. How effective are fiscal incentives for R&D? A review of the evidence. *Research Policy* 29: 449–469.

Healy, P. M., and J. M. Wahlen. 1999. A review of the earnings management literature and its implications for standard setting. *Accounting Horizons* 13 (4): 365–383.

Klassen, K., J. Pittman, and M. Reed. 2004. A cross-national comparison of R&D expenditure decisions: Tax incentives and financial constraints. *Contemporary Accounting Research* 21: 639–684.

Manzon, G. 1994. The role of taxes in early debt retirement. *The Journal of the American Taxation Association* 16: 87–100.

McCutchen, W. 1993. Estimating the impact of the R&D tax credit on strategic groups in the pharmaceutical industry. *Research Policy* 22: 337–351.

Myers, S., and N. Majluf. 1984. Corporate financing and investment decisions when firms have information that investors do not have. *Journal of Financial Economics* 13: 187–221.

Newey, W., and K. West. 1987. A simple positive, semi-definite, heteroscedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55: 703–708.

Phillips, J., M. Pincus, and S. Rego. 2003. Earnings management: New evidence based on deferred tax expense. *The Accounting Review* 78 (2): 491–521.

Rogers, W. 1993. Regression standard errors in clustered samples. *Stata Technical Bulletin No. 13.* College Station, TX: Stata Press.

Roychowdhury, S. 2006. Earnings management through real activities manipulation. *Journal of Accounting and Economics* 42 (3): 335–370.

Ryan, H., and R. Wiggins. 2002. The interactions between R&D investment decisions and compensation policy. *Financial Management* 31: 5–29.

Shevlin, T. 1990. Estimating corporate marginal tax rates with asymmetric tax treatment of gains and losses. *The Journal of the American Taxation Association* 11: 51–67.

Swenson, C. W. 1992. Some tests of the incentive effects of the research and experimentation tax credit. *Journal of Public Economics* 49: 203–218.

Tillinger, J. 1991. An analysis of the effectiveness of the research and experimentation tax credit in a q model of valuation. *The Journal of the American Taxation Association* 12: 1–29.

Yermack, D. 1995. Do corporations award CEO stock options effectively? *Journal of Financial Economics* 34: 237–269.

Zang, A. 2005. Evidence on the trade-off between real manipulation and accrual manipulation. Working paper, University of Rochester.

THE ACCOUNTING REVIEW
Vol. 83, No. 3
2008
pp. 735–756

On the Determinants of Measurement Error in Time-Driven Costing

Eddy Cardinaels
Tilburg University

Eva Labro
London School of Economics

ABSTRACT: Although employees' time estimates are used extensively for costing purposes, they are prone to measurement error. In an experimental setting, we research how measurement error in time estimates varies with: (1) the level of aggregation in the definition of costing system activities (aggregated or disaggregated); (2) task coherence (the extent to which the activities that require time estimates present themselves coherently or incoherently); and (3) when notice is given that time estimates will be required (in advance or after the fact), that is, whether participants know that time estimates will be required before they perform the activities. We also test on response mode (estimates in percentages or absolute time units). The results suggest an important trade-off between the level of aggregation and measurement error: increasing aggregation in the definition of activities leads to lower measurement error. Also, advance notification reduces measurement error, especially in settings with aggregated activities or incoherent tasks. Finally, we find a strong overestimation bias when participants provide time estimates in minutes, which may be problematic for Time-Driven Activity-Based Costing that advocates the use of estimates in minutes. These results are relevant to accountants and decision makers who want to assess and control the measurement error in their costing system and to professionals in related areas that make use of time estimates (e.g., billing, tendering).

Keywords: *costing system design; measurement and aggregation error; duration drivers; time-driven activity based costing.*

JEL Classification: *M41.*

Data Availability: *Experimental data can be obtained from the authors on request.*

The authors thank Marlys Lipe (editor), two anonymous referees, Gary Cokins, Andrea Drake, Anne Farrell, Sue Haka, Thomas Hemmer, Joan Luft, Ken Merchant, Wim Van der Stede, Laurence van Lent, Timothy West, and participants at workshops at the London School of Economics and Tilburg University, the EAA Conference in Gothenburg, the MAS Section Midyear Meeting in Tampa, SAS ABM Meeting Heidelberg, and the GMARS Conference in Copenhagen for comments. All errors are exclusively ours. The support of CIMA is gratefully acknowledged.

Editor's note: This paper was accepted by Marlys Lipe.

Submitted February 2006
Accepted November 2007

I. INTRODUCTION

Surveys on how much time employees spend on a variety of activities are a fixture of many costing systems (e.g., Goebel et al. 1998; Ittner 1999; Maher and Marais 1998; Major and Hopper 2005; Ellis-Newman and Robinson 1998; Narayanan and Sarkar 2002; Kobeissi et al. 1998). Time estimates are an important category of duration drivers, with many management accounting textbooks advocating their use for allocating resource costs to activities (e.g., Cooper and Kaplan 1998; Kaplan and Atkinson 1998). As the service industry grows in importance, the use of such duration drivers is expected to increase even more. Given this trend, a new variant of Activity-Based Costing (ABC), namely, Time-Driven ABC (Kaplan and Anderson 2004) has emerged that utilizes these estimates as the sole type of cost driver. However, in spite of the widespread use of time estimates, there has been little discussion about the error they may introduce. Experience indicates that employees' time estimates on the activities they perform are often inaccurate (Ittner 1999). It is therefore important to understand the sensitivity of costing systems to this error, as decisions based on erroneous costing figures are likely to be distorted.

Although theory suggests that cost data are rarely error free and that the conditions under which costing systems provide accurate costs are stringent (Noreen 1991; Datar and Gupta 1994), to our knowledge we are the first to empirically study measurement error on duration drivers. The difficulty of doing research in this area is the unavailability of an observable true cost benchmark that can be used to assess the error in a costing system. In this paper we use a unique experimental design that overcomes this difficulty: (1) participants perform a series of activities for which they subsequently provide time estimates, and (2) a computer accurately measures the time spent on each activity, providing a true benchmark against which the participants' time estimates are compared.

Our objective is to examine how measurement error in time estimates varies with prevalent costing system design and task parameters. First, we study the impact of the level of aggregation in the definition of a costing system's activities on measurement error, since aggregation is one of the most important aspects of accounting information. In the theoretical accounting literature, Datar and Gupta (1994) and Christensen and Demski (1995) highlight a potential trade-off between aggregation and measurement error; however, to our knowledge this trade-off has not previously been empirically explored. Second, we test whether the timing of the notification that time estimates will be required affects measurement error. Costing system design practices vary widely in this respect (Narayanan and Sarkar 2002; Major and Hopper 2005). Sometimes employees are notified that time estimates will be required before they perform their basic tasks (prospective notification), and on other occasions, typically with the installation of a new costing system, such notification is given after completion of the basic tasks (retrospective notification). Third, we explore the impact of task coherence on measurement error since the activities that require time estimates can present themselves coherently (in a logical structure) or incoherently (in a random order). Finally, since time estimates are given either in percentages or in absolute time units (Kaplan and Anderson 2004), we also conduct tests using both types of response mode.

We find support for the hypothesis that increasing aggregation leads to a decrease in measurement error (Datar and Gupta 1994; Christensen and Demski 1995). While ABC advocates argue that reducing the level of aggregation will lead to more accurate cost figures, our results suggest that such reduced aggregation goes hand in hand with increased measurement error in time estimates. As a consequence, an increase in overall accuracy from reduced aggregation cannot be guaranteed.[1] We also find that participants generate lower measurement error when they receive prospective notification, consistent with more cognitive resources being allocated to the time estimation task.

Additionally, our study provides evidence on the interactions of notification with aggregation and task coherence. Prospective notification is more effective in reducing measurement error when activities are defined on an aggregated level. However, when aggregation decreases, notification has no effect as the task of providing time estimates becomes highly complex. Prospective notification helps reducing measurement error when employees perform incoherent tasks, whereas it does not help when tasks are coherent as unconscious memory processes already do a good job at keeping track of time in the latter case. Finally, we find that when participants provide estimates in number of minutes, they strongly overestimate their time spent, which may be problematic in the application of the new Time-Driven ABC that advocates the use of minutes' estimates (Kaplan and Anderson 2004). This overestimation bias is likely to undo the intended effect of the minutes-based response on allocating only practical capacity rather than full capacity.

The contribution of our paper is fivefold. First, in dealing with the need to have a true benchmark, our experiment addresses several research questions on costing system design. While previous experiments document the impact of various (post-design) costing methods on decision-making (e.g., Callahan and Gabriel 1998; Gupta and King 1997), it is important to understand the factors that affect measurement error early in the design stage. Not controlling such error may result in a more erroneous costing system than intended. Second, we provide and examine two tools to control measurement error and to improve the robustness of a costing system to this error: the timing of notification and the level of aggregation. Third, we study the effect of task coherence on measurement error. While task coherence may not be directly under the control of the costing system designer, it is important to consider the impact it may have on measurement error when developing a costing system. Fourth, to the best of our knowledge, this study is the first to examine the effect on measurement error of both response modes used in practice: minutes and percentage estimates.

Fifth, we provide our participants with an incentive to make the most accurate estimates they can, while ruling out any error that could result from gaming behavior. As such, we establish a benchmark of the amount of error that arises when accuracy is the objective. This accuracy focus, which is typical for an accounting context, differs substantially from the cognitive psychology literature on time perception where such accuracy objective is absent. Although we partially rely on this literature to predict the impact of prevalent costing system design and task parameters, various other aspects of our accounting context differ substantially from the context and tasks employed in that literature. We introduce factors like aggregation and the above response modes that are not investigated in time perception studies but are of importance in a costing context. Also unlike time perception studies (Brown 1985), we combine prospective and retrospective notification in one study to enable us to assess the relevance of notification for controlling measurement error in costing system design.[2]

---

[1] Whether overall accuracy will increase or decrease will depend on the size of the trade-off between the two errors in a particular instance.

Our results have implications for practice. Specifically, we provide guidance on how to assess the measurement error that is likely to be introduced in costing figures, and on how to alleviate measurement errors in the design of a costing system. Besides costing applications, our insights are also relevant to any other area that utilizes time estimates. The remainder of the paper is structured as follows. Section II develops the hypotheses. Section III describes our experimental task setting and Section IV presents our results. Finally, Section V concludes and suggests avenues for future research.

II. HYPOTHESES

In this section, we develop our predictions of the impact of aggregation (H1) and task coherence (H2) followed by timing of notification (H3a), which can be before or after the basic task. We also explore the differential impact of such notification for the levels of task coherence and aggregation (H3b and H3c). Our hypotheses are tested for time estimates in percentages, but the "Results" section explores whether the findings are robust for a response mode in absolute minutes.

Aggregation

The level of aggregation in a costing system is a central design parameter. The costing system can be designed in an aggregated way where many activities are combined into few activities as defined in the costing system. Alternatively, the costing system can define many activities, and make less use of aggregation. The theoretical accounting literature has established that costing systems are unlikely to provide error-free costs (e.g., Noreen 1991; Datar and Gupta 1994). Datar and Gupta (1994) classify the possible errors as follows: "specification errors," which arise from using the wrong cost driver, "aggregation errors," which stem from adding together heterogeneous resources into cost pools, and "measurement error," which results from practical difficulty in identifying costs with a particular cost pool, or in measuring the specific units of the resources consumed by various activities.

Costing systems such as ABC are often devised in an effort to increase accuracy by reducing specification and aggregation error (e.g., Cooper and Kaplan 1988). However, Datar and Gupta (1994) show that partially refining a costing system by reducing specification and/or aggregation errors may well increase overall costing error due to the loss of partially offsetting effects between these errors. They indicate that a similar trade-off may exist between aggregation error and measurement error. For instance, reducing aggregation may reduce overall accuracy when the firm measures overhead at the aggregate cost pool fairly accurately, but significant error is introduced in measuring overhead at the disaggregated cost pools.

---

[2] Further, we use a clerical task that is more relevant to a real-life job context than the tasks used in the cognitive psychology literature: listening to music (Jones et al. 1993; Boltz 1998; Brown and Boltz 2002), tracing a figure (Brown 1985), sorting cards (Zakay and Shub 1998), visual search for a target letter (Brown 1997), or even sitting around idle (Brown and Boltz 2002). The tasks in the latter studies generally take only a few seconds (Block 1978; Brown 1985, 1997; Zakay and Feldman 1993; Brown and Boltz 2002) or even milliseconds (Allan 1979; Jones et al. 1993; Boltz 1998), whereas in a costing context task durations tend to be longer. The cognitive psychology literature itself questions whether results derived for such short durations will extend to longer durations (Francis-Smythe and Robertson 1999, 336). Since costing is intended to support business, we also explicitly explore performance in the main task, which is an important aspect of an employee's job, whereas most time perception studies only focus on the time estimation task (Brown 1997, 1120).

Datar and Gupta (1994, 569) refer to the use of time estimates as a source of measurement error. Relatedly, Christensen and Demski (1995) conjecture that the use of multiple cost pools (i.e., less aggregation) may not necessarily lead to more accurate product costs as measurement error may increase. When applied to measurement error on duration drivers, the argument is that measurement error increases with the number of activities for which the estimator has to make a time estimate.[3] For example, as Cooper et al. (1992, 12) note:

> A setup person might be able to specify quite accurately that setting up to produce a given product takes about 2.5 hours but be unable to break down the 2.5-hour elapsed time into the time required to perform 20 separate micro activities.

Although it might seem plausible that decreasing aggregation would lead to increased measurement error, to our knowledge no empirical costing studies have provided evidence on this relation.[4] Moreover, although it has been shown that splitting an event category into smaller subcategories can decrease accurate *recall of frequency* of events (Fiedler and Armbruster 1994), the cognitive psychology literature on duration perception has not explored this aggregation issue either, as all experiments require time estimates of one event.[5] We expect that reducing aggregation in the definition of activities leads to increased task complexity in the time estimation. The literature on task complexity suggests that an increase in the number of input cues (activities) (1) requires additional mental processing by a working memory with limited capacity (Miller 1956; Kahneman 1973; Drewnowski 1980) and (2) decreases clarity of processing (Bonner 1994). In turn, such increased task complexity has been shown to reduce judgment performance (Bonner and Sprinkle 2002). We thus hypothesize that:

H1: As aggregation in the definition of activities increases, measurement error in time estimates decreases.

Task Coherence

Tasks can be coherent or incoherent. When the task is coherent, the activities are presented and performed in a structured and systematic sequence. When a task is incoherent, activities are addressed as they come in. For example, a laborer may work at a particular task on a particular machine for several hours and only move on to another activity after his lunch break (coherent), while a secretary may take phone calls, type an invoice, reply to a manager's query, continue to type the invoice, etc. (incoherent). Another example is an assembly line (coherent) versus a flexible manufacturing cell (incoherent). Although the costing system designer does not determine task coherence, as it is driven by the underlying production technology, we will argue below that task coherence may have an important effect on measurement error. Also, given knowledge of whether the job design is coherent or incoherent, a decision maker using costing system data will be able to assess the impact of task coherence on the robustness of the costing system to measurement error.

---

[3] Note that we model aggregation at the activity cost pool level (following, e.g., Noreen 1991; Babad and Balachandran 1993; Datar and Gupta 1994) and therefore implicitly study measurement error on resource duration drivers (first stage of a cost allocation system). However, our results are also relevant to a second-stage allocation, where time estimates can be important when intensity drivers are used (Kaplan and Atkinson 1998).

[4] In accounting, Lim and Sunder (1990, 1991) study the trade-off between measurement error and the number of price indices used in asset valuation. The theoretical econometrics literature (Grunfeld and Griliches 1960) looks at situations in which aggregation reduces specification errors made at the micro equations level, but none of the aggregation gain situations they discuss are related to decreasing measurement error. They only mention in passing that poor-quality micro data may be another source of aggregation gain (Grunfeld and Griliches 1960, 10). Aigner and Goldfeld (1974) discuss econometric estimation and prediction using aggregate data when these are measured more accurately than their components, without touching on whether aggregate data are indeed measured more accurately.

[5] Brown et al. (1992) ask their subjects to attend to duration of multiple stimuli, but still only ask a time estimate for one of the stimuli. They research how the presence of *concurrent* stimuli impacts on the accuracy of this one estimate.

A line of research in cognitive psychology examines how the duration of events is encoded and retrieved from memory (Boltz 1998; Brown 1985; Brown and Boltz 2002). The primary assumption in this literature is that when individuals experience an event, their attention is split between the event's temporal and nontemporal information. Temporal information concerns time-related aspects of the event such as duration and frequency, whereas nontemporal information relates to the content of the event itself (Boltz 1998). Coherent events require less cognitive effort to encode than incoherent events, as the former contain an intrinsic hierarchical organization or structure. Thus, people require fewer resources to encode coherent events in memory (Brown and Boltz 2002), in which case more resources can be directed toward temporal information processing (keeping track of time). When nontemporal task demands increase due to the incoherence of the task, temporal processing receives less attention and, hence, duration judgments become less accurate (Brown 1985). As a consequence, time judgments should be more accurate for coherent than incoherent events. Boltz (1998), Jones et al. (1993), and Brown and Boltz (2002) find support for this prediction in music cognition experiments with durations of a few seconds, manipulating both coherent (musical melody) and incoherent events (arbitrary tones and rhythms).

Another strand in this literature suggests that temporal and nontemporal information is *jointly* encoded in memory. This literature is based on Ornstein's (1969) storage size hypothesis, which posits that an increasing number of stimulus events should increase storage size used in memory and lengthen the perceived experience of the events' duration. Increasing the number of stimuli via an incoherent task sequence should increase subjective duration. Several experiments show that greater randomness in stimulus sequences lengthens time judgments as participants rely on the number of stimuli within an array of nontemporal information to derive their estimates of temporal information (Ornstein 1969; Jones et al. 1993; Block 1985). However, other research in this strand of literature finds mixed results (e.g., Poynter 1983).

In sum, both strands in this literature predict less accuracy in time perception when events are presented incoherently rather than coherently (albeit with not fully conclusive evidence), but attribute this effect to different explanations. Also, it clearly remains an open question whether these prior findings extend to tasks that are relevant to a professional context and that typically require longer time intervals. Nevertheless, we predict:

> **H2:** As coherence in the (nontemporal) task increases, measurement error in time estimates decreases.

Notification of Time Estimation Task

The notification that participants will need to provide time estimates after performing a set of activities can be issued prospectively (before they start on the primary set of activities) or retrospectively (after they have finished the primary set of activities). For example, in the case of recurring revisions of costing systems, employees are notified that they will have to provide time estimates in advance (Ellis-Newman and Robinson 1998; Ittner 1999; Major and Hopper 2005), whereas when a new information system is installed or when consultants decide to update existing costing data, employees may be asked to provide time estimates after the fact (e.g., Narayanan and Sarkar 2002). Since costing system designers can choose the timing of notification, it is interesting to examine the consequences of this decision.

Boltz (1998) argues that prospective notification triggers the allocation of more mental resources toward cognitive timing. This should lead to more accurate and reliable estimates of the task's duration as the encoding of temporal information improves. However, this theory also argues that this may come at the expense of relatively poor performance on the task involving nontemporal information (here, slower execution of the main tasks performed by employees).

Conversely, retrospective notification directs the mental resources toward the processing of nontemporal information. Remembered durations will be based on unconscious memory-related processes and, hence, time estimates will have to rely on other available cues (Zakay and Feldman 1993). Remembered duration judgments are expected to be less accurate and more variable (Brown 1985; Boltz 1998). We predict:

H3a: When notification of the time estimation task is prospective as opposed to retrospective, measurement error in time estimates is decreased.

We are also the first to examine the interactions of notification and task coherence, and notification and aggregation. As such, we explore when costing system designers can expect prospective notification to have more incremental impact on reductions in measurement error. The cognitive psychology literature suggests that if no alternative mechanism to track time is available, as is the case for the retrospectively notified group, one resorts to unconscious memory processes to reconstruct time durations. In the case of incoherent tasks, this unconscious timer does not work very well when the basic task has many chunks or stimuli changes (Block 1985). Memory processes are disturbed by the high amount of contextual change (Block and Zakay 1997) and the number of subjective time units stored in memory will increase due to the many changes in stimuli (Ornstein 1969). Therefore, we expect that retrospective time estimation is more difficult and less accurate in incoherent than in coherent tasks.

In contrast, when participants are notified in advance that their time estimates will be required, they can consciously employ alternative ways of keeping track of time. Therefore, they are not as heavily affected by the disturbances of the memory process (Migliore et al. 2000) resulting from an increase in stimuli changes (Block 1992). Thus, prospective notification may be of greater benefit in an incoherent task, where the unconscious memory processes do not perform well. Accordingly, we predict:

H3b: The negative relation between timing of the notification and measurement error in time estimates (prospective notification resulting in less error than retrospective notification) is stronger when task coherence decreases.

The conscious cognitive timer that is used in settings where participants are prospectively notified, however, has its limitations. Miller (1956) and Drewnowski (1980) have shown that the conscious memory span for both encoding and retrieving is limited and decreases rapidly after around four items. Reducing aggregation increases the number of items to be encoded in and retrieved from memory. Therefore, reducing aggregation may increase the complexity of the temporal task to such extent that even a conscious timer cannot deal with it very well. This leads to the prediction that prospective notification helps improve performance of the time estimation task in an aggregated setting with a limited number of items, but the effect is lower when aggregation is reduced because of the limitations of the memory span.

> **H3c:** The negative relation between timing of the notification and measurement error in time estimates (prospective notification resulting in less error than retrospective notification) is stronger when aggregation increases.

III. EXPERIMENTAL PROCEDURES

Experimental Task

We employ a unique design to measure the accuracy of time estimates on a set of predefined activities performed by participants. Participants assumed the role of a worker in a firm producing customized designs of multi-colored wooden blocks. Each participant performed a set of 20 simple production and clerical tasks (as defined by Bonner et al. [2000]) related to activities such as machine set-ups, price quotes, and material requisitions. Participants were also required to provide time estimates on these performed activities. Before describing our experimental manipulations, participants and procedures, we briefly sketch the task setting and the set of activities, as shown in Figure 1.

In activity A "machine set-ups" were programmed and transmitted to production machines.[6] The task involved two subsets of programming activities, where the activity depended on whether the design ordered is "L-shaped formats" (A1) or "special-shaped formats" (A2). Participants typed a program code based on the customer order. The code required a statement for the type of program (activity A1 requires program A, whereas A2 requires program B), the number of units produced, and paint and location codes based on a 3×3 matrix.[7]

Activity B involved "price quotes." This activity was split into two further subsets of activities, in which quotes for a "single product" (B1) or "multiple products" (B2) were required. This was also an easy task that involved completing and submitting a standardized form to a client on the basis of incoming mail. The standard form, as displayed in Figure 1, required no calculations but involved filling out the client's name and email address as given in the incoming mail, the number of cuts and colors required for the design, and whether a special-shape surcharge (applicable if the design is not L-shaped) and a discount (if a client orders more than 500 pieces) applied.

Finally, in activity C "material requests" were prepared with requests for "wooden blocks" (C1) and "paint" (C2) as sub-activities. Based on supply logs received from workstations that require wood and paint, participants prepared and submitted a standard internal message to the materials-handling department. Participants entered the workstation in need of supply, and the amount of stock needed for replenishment in the case of paint supply or the amount of scrap removal and blocks needed in the case of wooden block requests.

Experimental Factors

All subjects first perform the basic (nontemporal) task, which consists of the 20 activities. They then provide estimates on the time spent on the activities, that is, they perform

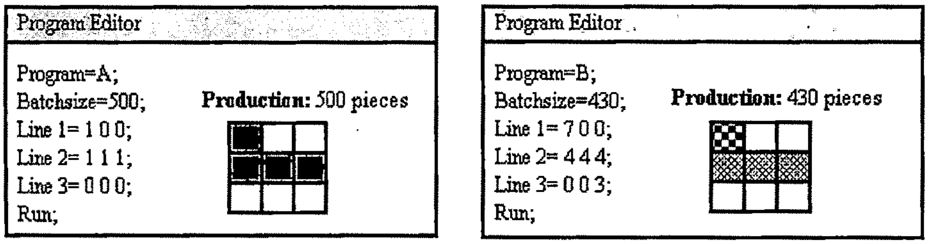
---

[6] The names of activities between quotation marks are those provided to the participants.

[7] A palette with nine colors is issued (not displayed here) providing color codes. By using different fills attached to a color, people suffering from color blindness could also assess these codes (http://www.vischeck.com). Codes 0, 1, 2, 3, 4, 5, 6, 7, 8, and 9 represent blank, red, orange, yellow, purple, brown, green, blue, cyan, and black, respectively. We also employ a gray background screen for the entire experiment, as gray has been shown to be the most neutral color in time perception studies (Gorn et al. 2004).

FIGURE 1
Screenshots of the Type of Activities Used In the Experimental Task[a]

Task A: Machine set-ups for L-shaped format (A1) and special shaped format (A2)



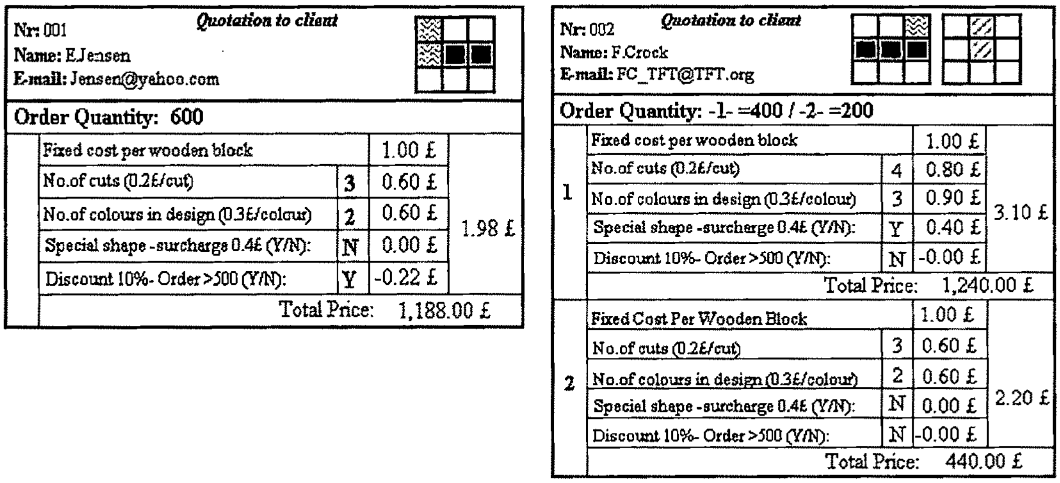
Program=A;
Batchsize=500;
Line 1= 1 0 0;
Line 2= 1 1 1;
Line 3= 0 0 0;
Run;

Production: 500 pieces

Program=B;
Batchsize=430;
Line 1= 7 0 0;
Line 2= 4 4 4;
Line 3= 0 0 3;
Run;

Production: 430 pieces

Task B: Price quotations for one product (B1) and multiple products (B2)



Quotation to client

Nr: 001
Name: E.Jensen
E-mail: Jensen@yahoo.com

Order Quantity: 600

Fixed cost per wooden block		1.00 £	1.98 £
No.of cuts (0.2£/cut)	3	0.60 £	
No.of colours in design (0.3£/colour)	2	0.60 £	
Special shape -surcharge 0.4£ (Y/N):	N	0.00 £	
Discount 10%- Order >500 (Y/N):	Y	-0.22 £	
Total Price:		1,188.00 £	

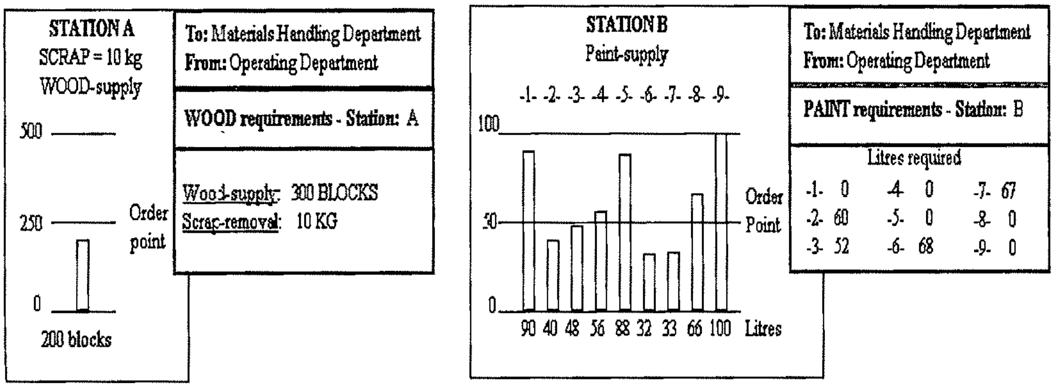
Quotation to client

Nr: 002
Name: F.Crock
E-mail: FC_TFT@TFT.org

Order Quantity: -1- =400 / -2- =200

1	Fixed cost per wooden block		1.00 £	3.10 £
	No.of cuts (0.2£/cut)	4	0.80 £	
	No.of colours in design (0.3£/colour)	3	0.90 £	
	Special shape -surcharge 0.4£ (Y/N):	Y	0.40 £	
	Discount 10%- Order >500 (Y/N):	N	-0.00 £	
	Total Price:		1,240.00 £	
2	Fixed Cost Per Wooden Block		1.00 £	2.20 £
	No.of cuts (0.2£/cut)	3	0.60 £	
	No.of colours in design (0.3£/colour)	2	0.60 £	
	Special shape -surcharge 0.4£ (Y/N):	N	0.00 £	
	Discount 10%- Order >500 (Y/N):	N	-0.00 £	
	Total Price:		440.00 £	

Task C: Material requests for wood (C1) and paint (C2)



To: Materials Handling Department
From: Operating Department

WOOD requirements - Station: A

Wood-supply: 300 BLOCKS
Scrap-removal: 10 KG

To: Materials Handling Department
From: Operating Department

PAINT requirements - Station: B

Litres required

-1-	0	-4-	0	-7-	67
-2-	60	-5-	0	-8-	0
-3-	52	-6-	68	-9-	0

---

[a] All participants perform the same 20 task stimuli: three times A2, B1, C1, C2, and four times A1 and B2.

the time estimation (temporal) task. We employ a 2 × 2 × 2 between-subjects design, in which factors of interest are manipulated in the manner as displayed in Panel A of Table 1. The first factor is the task coherence of the basic task. The 20 task activities can be presented coherently, when participants first perform machine set-ups (A1, then A2), followed by price quotes (B1, then B2), and material requisitions (C1, then C2). In incoherent tasks these activities are introduced in no particular order, as displayed in the footnote to Table 1.

TABLE 1

Description of the Experimental Factors and Definition of the Dependent Variable

Panel A: Manipulated Between-Subject Factors on the Experimental Timeline[a]

	Basic Task		**Time Estimation Task**	***Time in Experiment***
NOTIFICATION	**TASK COHERENCE**	**NOTIFICATION**	**AGGREGATION**	
when Prospective	Coherent: Tasks of Figure 1 appear in a sequential order Incoherent: Tasks of Figure 1 appear in a random order	when Retrospective	Three: participants estimate time spent on the three main activities A, B and C Six: participants estimate time spent on the six sub-activities A1, A2, B1, B2, C1 and C2	

Panel B: Dependent Variable: Measurement Error in the Percentage Estimation

Formula for the measurement error in the percentage estimation[b]

Requirement: divide 100% among the set of main/sub-activities

$$\text{Total measurement error} = \Sigma\ |\%\ \text{actual time}_{\text{main/sub-activity}} - \%\ \text{time estimated}_{\text{main/sub-activity}}|$$

Example of Comparability of This Metric across the Levels of Aggregation

Aggregation = 3	**A**	**B**	**C**	
% actual time	30	45	25	100
% time estimated	20	50	30	100
Total measurement error (%)	10	5	5	20

Aggregation = 6	**A1**	**A2**	**B1**	**B2**	**C1**	**C2**	
% actual time	17	13	25	20	10	15	100
% time estimated	10	10	25	25	15	15	100
Total measurement error (%)	7	3	0	5	5	0	20

---

[a] CAPITALS denote the three factors; *underlining* denotes the levels of the factors as manipulated. Labels A, B, and C refer to the main activities machine set-ups (A), price quotations (B) and material requests (C) with their sub activities L-shaped (A1) versus special-shaped machine set-ups (A2), one (B1) versus two product price quotes (B2) and material requests for wood (C1) versus paint (C2); as described in Figure 1. Note that for coherent tasks the 20-task stimuli appeared as follows: A1-A1-A1-A1-A2-A2-A2-B1-B1-B1-B2-B2-B2-B2-C1-C1-C1-C2-C2-C2. Incoherent tasks appeared as follows: A1-B2-C2-B1-C1-A2-B2-A1-B1-A2-C2-B2-C1-B1-C2-A1-C1-B2-A2-A1.

[b] Summary statistics of the dependent variable, averaged across all cells: Mean (median): 31.95 (30.89), Min: 2.78, Max: 72.18, SD: 13.32. Note that lower values represent lower measurement error.

The second factor is the level of aggregation. In the time estimation task, we vary the number of activities participants have to assess in terms of time consumption. More aggregation requires estimates on the three main activities (A, B, C), whereas less aggregation requires estimates on the six subsets of activities (A1 through to C2).[8] The third factor we vary is the timing of the notification that time estimates will be required. Some participants receive prospective notification; they know that time estimates will be asked of them before they start their basic tasks. Others receive no notification before they start their basic tasks. They are retrospectively notified about the fact that time estimates are required.

Dependent Variable

Based on the participants' percentage estimates of time, we calculate total measurement error as reported in Panel B of Table 1, to use in our hypotheses tests. Total measurement error is defined as the sum of the absolute deviations of the estimated time spent on an activity versus the actual time for that activity. We report total error rather than mean error (as, for example, in Datar and Gupta [1994] or Babad and Balachandran [1993])[9] to assure that the measurement error is comparable across the two levels of aggregation as Panel B of Table 1 shows.

In practice two response modes are common when people are surveyed about the time they spend on various activities: percentages and absolute time units (Kaplan and Atkinson 1998). Sometimes these are offered in combination (Kobeissi et al. 1998; Kaplan and Atkinson 1998), as, for example, in Prodacapo's Activity Tracker. Because we do not have any reason to expect the accuracy of these response modes to be significantly different, we introduce a within-subject variation of this factor to test the robustness of our results. Specifically, after providing a percentage estimation of the time spent on each activity, participants provided their estimates in minutes via an estimate of the total amount of minutes worked, which they subsequently divided over the various activities.[10] To verify the robustness, we recode the minutes-based estimates as percentages by dividing the absolute estimates for each activity by the total time estimate. Applying the formula of Panel B of Table 1 on these recoded figures, we obtain an alternative error for minutes' estimates that is comparable to the measurement error on percentage estimates.

Participants and Procedures

We recruited 157 undergraduate students taking a management accounting course at a large academic institution in the U.K. On average students were 20.4 years old and 90 percent of them reported having gained some business experience, with a mean of 8.5 months of experience acquired through internships or part-time jobs. We chose to use student subjects as workers because the experiment did not involve or require any specific accounting knowledge (Rose and Wolfe 2000) and the estimation tasks required only a limited level of skill (Libby and Lipe 1992; Bonner and Sprinkle 2002). Moreover, in practice, any member of an organization can be asked to provide time estimates, irrespective of their experience.

---

[8] This falls within the usual range of number of activities (two to nine) for which time estimates are requested of one staff member (Cooper et al. 1992) when designing a costing system. Of course, combined over all staff, the costing system can include many more activities.

[9] Datar and Gupta (1994) and Babad and Balachandran (1993) ensure comparability of mean error metrics by keeping the number of cost objects the same across their settings. Since we focus on activity level and compare a highly aggregated setting with a setting in which aggregation is reduced, we can only compare the means of the total error.

[10] This is consistent with what is done in practice when absolute time estimates are requested. A percentage-based response mode ensures that there is no measurement error on the total units of activity, an assumption that is also made by Datar and Gupta (1994). For an absolute response mode, measurement error on the total units of time can creep in as an additional problem, which we will explore in more detail later.

We issued experimental materials via a computer and randomly assigned participants to one of the experimental conditions upon entering the computer room. Sessions lasted about an hour.[11] The procedures were as follows. In the first part of the experiment, we collected the participant's demographic information. Next, we introduced the case company and described all activities in full. To get acquainted with the activities, participants performed trial examples of machine set-ups, price quotes, and material requisitions immediately after we described each activity. In the second part of the experiment, we issued the basic task with the 20-task stimuli. Participants could only move on to the next task after correctly completing the previous task. Therefore, speed of completion is our only measure of performance of the basic task. The computer accurately measured the time spent on each of the 20-task stimuli.[12]

Depending on their assigned condition, the notification of the time estimation task was presented immediately before or after these 20-task stimuli. In either case, we stressed that participants provide time estimates as accurately as possible, as the company required estimates for designing an accurate ABC system.[13] Participants were then required to provide the time estimates either on the three major activities or on the six sub-activities, first in percentages of time spent and next in number of minutes spent on each activity. Their confidence in both the percentage as well as the minutes-based estimation was self-assessed on a seven-point Likert scale. The experiment ended with a debriefing questionnaire containing items on task motivation, understanding, and realism, and manipulation checks regarding the case materials.[14]

At the start of the experiment, participants were informed that performance-based pay is involved. In order to ensure effort was exerted, participants received a flat wage of £3 ($6) for participation and a bonus of a maximum of £4 ($8) for their performance, as measured by both the speed with which they correctly performed the 20-task stimuli as well as the accuracy of their estimates in the time estimation task. This payoff structure was revealed only at the time the basic task or the time estimation task was described.[15] A debriefing item of self-perceived motivation (on a seven-point scale) suggests on average a high motivation (mean = 5.37), with no significant differences among cells (all p's > 0.17). The debriefing questionnaire further does not indicate significant differences among participants related to the time pressure experienced, the perceived purpose of the time estimates, and task understanding.

---

[11] To help students plan their schedules and commit to a time slot for using the computers, students were told that it was expected that the task would take slightly less than one hour. Specifying an expected duration may reduce variation in this measure, leading to a reduced likelihood of obtaining results that support variation in duration and estimated duration across the experimental groups (Libby and Lipe 1992).

[12] Each of the 20 task stimuli is provided on one screenshot. The computer measures accurately the time spent on each screen. The activities vary in length of time needed to execute them, representing a situation in which it would be beneficial to introduce an ABC system. Participants work without a time limit in order not to confound our results with the effects of time pressure that have been identified in the literature (e.g., Flückiger 1985; Zakay 1993; Ariely and Zakay 2001). We test whether participants felt stressed for time in our exit questionnaire. Both the model as well as the factors of interest are not significant (all p's > 0.23).

[13] In the exit questionnaire, participants in all cells indicate that they were clear that the data they provided were not used for any other purpose. In addition, as we repeatedly state that management requires accurate time estimates, they do not deliberately over- or underestimate the time spent on the specific activities as a result of gaming behavior (no differences across cells; ANOVA on item for overestimation, all p's > 0.13; ANOVA on item for underestimation, all p's > 0.34).

[14] Participants are not allowed to make notes or communicate. They are cautioned not to discuss the experiment with each other until all the sessions had been run, both for reasons of research validity and maximization of their own bonus payment. If participants were to discuss the experiment this would have weakened the power of the tests and the likelihood of finding the hypothesized effects especially with regard to the timing of notification effect.

Manipulation Checks

With regard to the factor aggregation, an item of the post-questionnaire confirms that the number of different activity categories specified under aggregation (three versus six) has an effect on the participants' perception of task categories. In particular, participants estimating the time spent over six activities indicated that they had more difficulty in apportioning time across task categories as a result of *too many task categories*[16] than participants estimating the time spent over three activities ($F_{aggr\ (1,149)} = 6.26$; $p < 0.02$; mean is 3.83 on seven-point Likert scale for three activities group and 4.54 for six activities group). Also the perception of task coherence is consistent with the way we manipulate this factor. Specifically, participants who are given a coherent task perceive it as following a more logical structure (mean is 5.42 on a seven-point Likert scale) than those assigned an incoherent task (mean is 4.40; $F_{task} = 22.55$; $p < 0.01$). Finally, in case of prospective notification, we clearly notify participants about the time estimation task before they start their basic task of 20 activities. Under retrospective notification, we only issue this notification after participants have performed these tasks.

IV. RESULTS

We first describe the findings of our hypotheses tests. Supplementary analyses then explore some additional effects of notification and test the robustness of our results across the two response modes. Hypotheses tests use a full-factor ANOVA model on the log of the measurement error of the percentage estimation. We use the log of errors to reduce the effect of outliers (Waller et al. 1999, 729). In addition, we also report an effect size measure for each hypothesis based on the specific contrast estimate within the ANOVA-analyses (Buckless and Ravenscroft 1990). Summary statistics of the measurement error and the log of this error per experimental cell are reported in Panel A of Table 2. Panel B of Table 2 reports the full factor ANOVA analyses. Panel C explores the interactions between notification and task coherence, and notification and aggregation in more detail.

Main Effects

Panel A of Table 2 suggests that measurement error in a costing system that exhibits high aggregation (as six activities are aggregated into three) is lower than under a costing system that is less aggregated (as six activities are defined). Panel B of Table 2 provides evidence of a strong effect of aggregation (1 percent significance level) on the log of the

---

[15] Following Vera-Muñoz (1998), we pay participants according to quartiles of ranked performance. For the basic task, the 25 percent fastest participants receive a bonus of £2, the following 25 percent £1.3, and the next 25 percent £0.7; the remaining subjects do not receive a bonus. For the time estimation task, we pay subjects according to quartiles of accuracy following an analogous structure (£0–2), using an average measure of accuracy in both percentage and minutes estimations. Hence, both tasks are equally rewarding. *De facto*, we employ the quartile divisions on speed and accuracy in every experimental cell to increase the *ex post* perception of fairness. However, we reveal this to the participants only subsequent to the experiment in a debriefing lecture.

[16] Participants were asked to indicate on a seven-point Likert scale ranging from completely disagree to completely agree to what extent they agreed with the statement "There were too many task categories making estimation of time allocated across tasks categories difficult."

TABLE 2
Summary Statistics and Hypotheses Tests

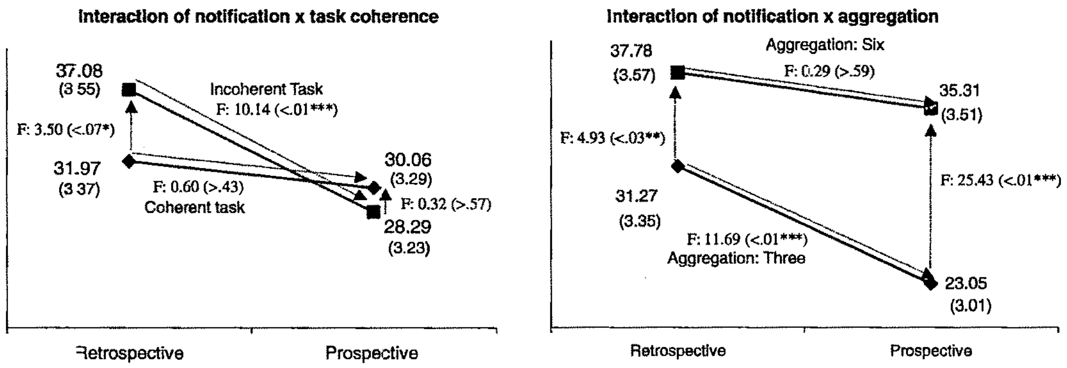
Panel A: Summary Statistics of the Measurement Error[a]

Retrospective Notification				Prospective Notification			
Aggregation	**Coherent Task**	**Incoherent Task**	**Total**	**Aggregation**	**Coherent Task**	**Incoherent Task**	**Total**
Three	30.62 (3.28) [n = 20]	31.92 (3.42) [n = 20]	31.27 (3.35)	Three	24.42 (3.05) [n = 19]	21.67 (2.97) [n = 19]	23.05 (3.01)
Six	33.32 (3.45) [n = 20]	42.24 (3.69) [n = 20]	37.78 (3.57)	Six	35.71 (3.53) [n = 20]	34.91 (3.50) [n = 19]	35.31 (3.51)
Total	31.97 (3.37)	37.08 (3.55)	34.53 (3.46)	Total	30.06 (3.29)	28.29 (3.23)	29.18 (3.26)

Panel B: Full Factor ANOVA on the Log of the Measurement Error and Hypotheses Tests

Source of Variation	Test	Df	SS	F-value	p-value
Aggregation (A)	H1	1	5.1509	26.58	0.0001***
Task coherence (T)	H2	1	0.1597	0.82	0.3655
Notification (N)	H3a	1	1.5249	7.87	0.0057***
Notification × Task Coherence (N × T)	H3b	1	0.5683	2.93	0.0889*
Notification × Aggregation (N × A)	H3c	1	0.8102	4.18	0.0426**
Aggregation × Task Coherence (A × T)	/	1	0.0652	0.34	0.5627
Notif. × Aggr. × Task Coh. (N × A × T)	/	1	0.0048	0.02	0.8750
Error		149	28.8754		

Panel C: Detailed Tests of the Interactions with Notification[b]


Interaction of notification x task coherence

Interaction of notification x aggregation

***, **, * Indicate significance levels of 1 percent, 5 percent, and 10 percent, respectively, (based on Type III sum of squares).

[a] This panel contains the means of the measurement error in the percentage estimation, (the log of this measurement error, which is used to test the hypotheses), and [the number of participants].

[b] The figure on the left-hand side explores the task coherence × notification interaction in detail. The effects of notification (task coherence) are explored under the two levels of task coherence (notification). The figure on the right, offers a similar analysis for the aggregation × notification interaction. Comparisons are made on the basis of contrast analyses (setting the cells not under consideration to zero). For each comparison (as indicated by arrows) we present the F-value and resulting p-value (two-tailed).

measurement error.[17] Consistent with H1, the effect size of aggregation is positive ($b = 0.362$; $t = 5.16$, $p < 0.01$, one-tailed), which means that a higher measurement error is made when six activities are defined than when three activities are defined. We are the first to empirically document that costing system designers can expect increases in measurement error when reducing the level of aggregation of cost pools. Attempts to improve costing accuracy by increasing the number of cost pools defined in a costing system should therefore be treated with caution.

Our ANOVA analysis does not support our hypothesis (H2) that task coherence has an effect on the log of the measurement error, as this factor is not significant in Panel B of Table 2. While the effect size is positive, it is not significant ($b = 0.064$; $t = 0.91$; $p = 0.18$, one-tailed), suggesting that incoherent tasks do not significantly increase the measurement error when compared to coherent tasks. The means reported in Panel A of Table 2 suggest that measurement error is generally lower when decision makers are prospectively notified about the time estimation task (H3a) than when they are retrospectively notified. This factor significantly affects the log of the measurement error in Panel B of Table 2. Indeed, the size of the effect suggests that prospective notification strongly decreases the measurement error ($b = -0.197$; $t = -2.81$, $p < 0.01$, one-tailed). This supports H3a. Presumably prospective notification facilitates the encoding of temporal information, because it allows participants to allocate mental resources to conscious cognitive timing.

Note that the factors "aggregation" and "notification" significantly affect participants' confidence in the accuracy of their percentage estimation (as participants indicated on a seven-point Likert scale after providing these estimates). Consistent with our predictions, participants presented with an aggregated set of activities were more confident in the accuracy of their estimates ($F_{aggr\ (1,149)} = 3.83$, $p < 0.06$; mean is 4.80 (resp. 4.48) for the group with three (resp. six) activities). Also, people that are prospectively notified have higher confidence in the accuracy of the time estimates they provide ($F_{notif.\ (1,149)} = 3.62$, $p < 0.06$; mean is 4.80 (resp. 4.49) under prospective (resp. retrospective) notification). Although we do not find support for our second hypothesis that task coherence decreases measurement error, an item in the post-questionnaire reveals that participants in incoherent tasks had to concentrate more fully when executing activities ($F_{task\ (1,149)} = 4.72$; $p < 0.04$; mean incoherent is 5.15; mean coherent is 4.63). Incoherent tasks seem to require more cognitive effort, as suggested in our hypothesis section. Further analysis of the interaction effects (in the next section) shows that task coherence will have an effect on measurement error if participants are retrospectively notified.

Interaction Effects

Panel B of Table 2 suggests that the interactive effect of notification by task coherence is significant. Consistent with H3b, the effect size of this interaction is negative and significant ($b = -0.241$, $t = -1.71$, $p = 0.04$, one tailed), indicating that prospective notification reduces the measurement error more strongly when tasks are incoherent. The results in Panel B further suggest a significant interaction of aggregation and notification. The size of this effect is positive and significant ($b = 0.287$, $t = 2.04$, $p = 0.02$, one tailed). As predicted in H3c, it suggests that the effect of prospective notification on reducing measurement error is smaller when aggregation decreases from three to six activities.[18] To further disentangle the effects of notification, we test the individual effects of notification on the log of the measurement error at each level of task coherence and aggregation (and also the effects of task coherence and aggregation at each level of notification), as plotted in Panel C of Table 2. In this panel, we again use contrast analyses as this allows us to make these specific comparisons, setting the contrast codes for the cells not under consideration equal to 0.

---

[17] Because of the hypothesized interaction effect with notification, we also check the main effect of aggregation for each level of notification. As we will show further in the section "interaction effects," aggregation is significant for both levels of notification (prospectively notified: $F = 25.43$; $p < 0.01$; retrospectively notified: $F = 4.93$, $p < 0.03$).

The results in Panel C of Table 2 suggest that prospective notification can reduce measurement error in the time estimates when the task is incoherent ($F = 10.14$, $p < 0.01$), but has no effect when tasks are coherent ($F = 0.60$, $p > 0.43$). Under retrospective notification, the only aid participants can rely on for reconstruction of time duration is their unconscious memory process. This works fairly well in coherent task settings, but fails in incoherent task settings, where the amount of stimuli and contextual changes disrupt its functioning. Therefore, prospective notification is of benefit when tasks are incoherent, but is irrelevant in a coherent task setting, where these unconscious memory processes work fairly well. Although we do not find a significant main effect of task coherence (H2), Panel C indicates that task coherence does affect those participants that are retrospectively notified ($F = 3.50$, $p < 0.07$). Here, incoherent tasks increase the measurement error. Task coherence is insignificant for those that are prospectively notified ($F = 0.32$, $p > 0.57$) as they presumably rely on a conscious cognitive timer, which is not affected by stimuli changes.

The results in Panel C of Table 2 further suggest that notifying the employee in advance decreases measurement error more strongly in aggregated settings ($F = 11.69$, $p < 0.01$). When aggregation in the definition of activities decreases (definition of six activities), the effect of notification is no longer significant ($F = 0.29$, $p > 0.59$). This is consistent with our hypothesis (H3c) that even a conscious cognitive timer does not work well under the increased complexity of a disaggregated setting, as memory span is too limited to keep track of six durations. Note that in contrast with task coherence that is only significant when people are retrospectively notified, aggregation has a stronger effect. It remains significant both when people are prospectively notified ($F = 25.43$; $p < 0.01$) as well as retrospectively notified ($F = 4.93$, $p < 0.03$).

Other Notification Effects

Given that notification has a strong impact on reducing measurement error, we explore further (1) the possibility of an intrusive notification effect in terms of reduced performance in the basic task, and (2) an additional motivation for a notification effect on measurement error through increased effort in time estimation. In the "Hypotheses" section we suggest a potentially intrusive effect of prospective notification: notified people may perform the basic (nontemporal) task more slowly as they devote resources to keeping track of the elapsed time. We measure performance in the basic task as the time to complete the total set of 20 activities. Panel A in Table 3 presents summary statistics for this variable. The full-factor ANOVA model with the dependent variable "performance in the basic task" is not significant ($F_{model} = 1.25$, $p > 0.27$), and is therefore not reported in Table 3. The effect of

---

[18] As mentioned earlier, we performed all hypotheses tests on the log of the measurement error to reduce the impact of outliers. However, the Grubb's test detected one significant outlier ($p < 0.05$) in our sample. All effects stay the same when the untransformed measurement errors are analyzed for the sample without the outlier. Analyzing the untransformed errors with the outlier included, only leads to a lower effect size for the interaction of aggregation × notification ($b = 5.75$, $t = 1.48$, $p = 0.07$, one-tailed). The effect size for the interaction of task coherence × notification ($b = -6.88$, $t = 1.77$, $p = 0.04$, one-tailed) is similar in terms of its p-value.

TABLE 3
Additional Effects of Notification

Panel A: Summary Statistics of Minutes Spent on the Basic Task and on the Time Estimation Task[a]

Retrospective Notification				Prospective Notification			
Aggregation	**Coherent Task**	**Incoherent Task**	**Total**	**Aggregation**	**Coherent Task**	**Incoherent Task**	**Total**
Three	19.32 {3.60} [n = 20]	18.02 {3.56} [n = 20]	18.67 {3.58}	Three	17.51 {3.73} [n = 19]	17.43 {4.04} [n = 19]	17.47 {3.88}
Six	16.23 {5.10} [n = 20]	18.38 {5.20} [n = 20]	17.30 {5.15}	Six	18.52 {6.90} [n = 20]	17.75 {7.15} [n = 19]	18.13 {7.02}
Total	17.78 {4.35}	18.20 {4.38}	17.99 {4.36}	Total	18.01 {5.32}	17.59 {5.59}	17.80 {5.45}

Panel B: Full Factor ANOVA of Effort in the Time Estimation Task

Source of Variation	Df	SS	F-value	p-value
Aggregation (A)	1	217.4809	37.77	0.0001***
Task coherence (T)	1	0.9047	0.16	0.6924
Notification (N)	1	46.5678	8.09	0.0051***
Notification × Task Coherence (N × T)	1	0.6182	0.11	0.7436
Notification × Aggregation (N × A)	1	24.1754	4.20	0.0422**
Aggregation × Task Coherence (A × T)	1	0.0202	0.00	0.9529
Notif. × Aggr. × Task Coh. (N × A × T)	1	0.1033	0.02	0.8936
Error	149	858.0306		

***, ** Indicates significance level of 1 percent and 5 percent, respectively, (based on Type III sum of squares).

[a] The table respectively contains means of minutes spent on the basic task, {minutes spent on the estimation task}, and the [number of participants].

notification is also not significant ($F_{notif.\ (1,149)} = 0.10$, $p > 0.76$), which implies that prospectively notified participants perform the basic task as fast as those that are retrospectively notified. We do not find an intrusive notification effect on basic task performance in our setting.

A second test explores an additional potential effect of prospective notification, besides the direct effect of more mental resources being allocated to cognitive processing of elapsed time. Specifically, participants who are prospectively notified might be spending more effort on the estimation task than those who are retrospectively notified. Following Libby and Lipe (1992), we proxy (duration of) effort by the time spent on the estimation task. Panels A and B of Table 3 show that more effort is triggered when participants are prospectively notified ($F_{notif.\ (1,149)} = 8.09$, $p < 0.01$).[19] This extra effort, induced by prospective notification, may additionally help to decrease measurement error on top of such notification allowing for more mental resources being allocated to the cognitive timer during encoding. Our initial results suggest that this may indeed be the case in an incoherent task setting.

---

[19] Note that reducing aggregation increases the time spent (see Table 3). This effect is not particularly interesting since reducing aggregation mechanically should lead to more time spent in estimation, as it takes longer to provide estimates for six than for three activities.

We compute partial correlations in an incoherent task setting, controlling for time differences as a result of aggregation. Both notification as well as effort in time estimation are negatively correlated with measurement error ($r = -0.384$, $p < 0.01$ and $r = -0.204$, $p < 0.08$, respectively).

Response Mode

A repeated measure test using the log of the estimation errors for the response modes minutes and percentages (correlation $r = 0.785$, $p < 0.01$) as a within-subjects factor indicates that our results are robust to response mode. In particular, there is no main effect of response mode on measurement error ($F_{\text{response }(1,149)} = 1.64$, $p = 0.203$), and there are no significant interactions with our factors of interest (all p's > 0.26). This suggests consistency across response modes. When we run the model using the minutes-based response mode separately, all previously identified effects remain significant, with the exception of the interaction effect of task coherence and notification (although means still go in the same direction). Given that in a "traditional" ABC system estimates in minutes are recalculated to percentages to allocate resource cost and that we do not find a significant main effect of response mode on measurement error, we conclude that in such costing systems, time estimates in minutes and in percentages are equally valuable in terms of accurately allocating the resource cost.

Two points warrant attention here. First, our test for self-perceived confidence across estimation response modes yields a significant effect of response mode ($F_{\text{response }(1,149)} = 15.39$, $p < 0.01$). The means suggest that participants have greater confidence in their percentage estimation (mean: 4.64 on a seven-point Likert scale) than in their minutes-based estimation (mean: 4.37). Second, part of the argument for the superiority of Time-Driven ABC over traditional ABC (Kaplan and Anderson 2004) is based on the idea that the minutes-based response mode is superior to the percentage-based response mode as the former allows that one allocate only the practical capacity cost, rather than the full capacity cost. Since percentages add up to 100 and few people report that a significant percentage of their time is idle, Kaplan and Anderson (2004) consider a direct estimate in minutes superior. Our results, however, suggest that the minutes-based response mode is associated with an important overestimation bias: more than 77 percent of our participants consistently overestimate the time spent on all activities. The average level of overestimation is about 6.64 minutes (about 37 percent), which is significantly different from zero ($t = 5.01$, $p < 0.01$), and is equally high for all experimental cells (all p's not significant). This overestimation bias is likely to undo the intended effect of the minutes-based response mode on the allocation of capacity in Time-Driven ABC.

V. CONCLUSIONS

We use an experimental setting to study the determinants of measurement error in time-driven costing. Our results are as follows. First, we find that there exists an important trade-off between aggregation and measurement error: reducing aggregation leads to increased measurement error. This trade-off needs to be carefully considered when making costing system design choices. Second, we show that prospective notification can reduce measurement error. This does not come at the cost of worse performance on the main task, and the effect is larger in more aggregate settings. When aggregation is reduced, the task of providing time estimates becomes increasingly complex due to a limited memory span. Therefore, even a conscious timer has difficulty to keep track of time in such setting.

Although we do not find a main effect of task coherence, it plays a role by interacting with notification; prospective notification reduces measurement error in the time estimates when tasks are incoherent, but not when tasks are coherent. We attribute this difference to the contextual and stimuli changes in incoherent settings that disturb the unconscious memory processes used when participants are retrospectively notified. Hence, prospective notification helps to reduce measurement error in incoherent settings. Conversely, unconscious memory processes work fairly well given coherent tasks, rendering prospective notification less important.

Third, we document that the minutes-based response mode is associated with a large overestimation bias, highlighting the need for caution in the use of Time-Driven ABC. Time-Driven ABC advocates argue that a minutes-based response mode is superior to one in percentages as the former requires that one allocate only the practical capacity cost, rather than the full capacity cost (Kaplan and Anderson 2004). Indeed, when submitting estimates in percentages, few employees tend to report a percentage of their time is spent idle. However, the presence of a large overestimation bias in the minutes' estimation is likely to undo the intended benefit of moving away from a percentage-based response mode.

Our results have direct implications for activity-based costing practice. Service-based economies rely heavily on time estimation in their cost allocations, with many areas such as billing, tendering, and project accounting tracking activities' time estimates. Given the extent to which time estimates are used, decision makers who use the figures produced by costing systems should be aware of the extent of measurement error that impacts the accuracy of the costing figures so that they avoid making flawed decisions. Our results also highlight the need for costing system designers to find ways to reduce the probability of such error in a costing system's design, and offer several tools that can be used toward this end. For example, as the highest error in time estimates is incurred when tasks are incoherent and activities are disaggregated, these cases would benefit most in terms of increased accuracy from an investment in an automatic on-line measurement system.

Our study suggests several avenues for future research. First, it would be worth exploring whether our results generalize to alternative tasks and experience levels. Following Brown (1985), who finds that inaccuracy in time estimates increases when the basic task becomes more demanding, we speculate that measurement error may become more severe when we use tasks that are more memory-processing-intensive than our simple production and clerical tasks. In contrast, measurement error is likely to decrease with experience in the basic task. Extensive practice permits some complex operations to occur with minimal mental capacity (Libby and Lipe 1992), freeing capacity for the time estimation task.[20]

Second, here, we deliberately abstract away from gaming in the experiment by rewarding participants on accuracy of their time estimates. Future research that includes the gaming behavior of participants when other objectives for the use of the cost data are introduced (such as that described in Merchant and Shields [1993], for example) can use our results as a gaming-free benchmark of achievable accuracy levels when maximum accuracy is the only underlying objective. Finally, the new Time-Driven ABC approach as advocated by Kaplan and Anderson (2004) suggests that estimates rely on informed managers' time estimates rather than on employee surveys. Anecdotal evidence by Cooper et al. (1992) suggests that managers also make errors in estimating the time that employees spend on various activities. It would be interesting to explore how the accuracy of managers' estimates compares to employees' estimates.

---

[20] However, experience in the time estimation task itself is unlikely to change our results. Hasher and Zacks (1979) show that memory processes that keep track of time are innate and therefore unlikely to record any developmental trends.

REFERENCES

Aigner, D. J., and S. M. Goldfeld. 1974. Estimation and prediction from aggregate data when aggregates are measured more accurately than their components. *Econometrica* 42 (1): 113–134.

Allan, L. G. 1979. The perception of time. *Perception and Psychophysics* 26 (5): 340–354.

Ariely, D., and D. Zakay. 2001. A timely account on the role of duration in decision-making. *Acta Psychologica* 108: 187–207.

Babad, Y. M., and B. V. Balachandran. 1993. Cost driver optimization in activity-based costing. *The Accounting Review* 68 (3): 563–575.

Block, R. A. 1978. Remembered duration: Effects of event and sequence complexity. *Memory & Cognition* 6 (3): 320–326.

———. 1985. Contextual coding in memory: Studies of remembered duration. In *Time, Mind, and Behavior*, edited by J. A. Michon, and J. L. Jackson. New York, NY: Springer.

———. 1992. Prospective and retrospective duration judgment: The role of information processing and memory. In *Time, Action and Cognition: Towards Bridging the Gap*, edited by F. Macar, W. J. Friedman, and V. Pouthas. New York, NY: Springer.

———, and D. Zakay. 1997. Prospective and retrospective duration judgments: A meta-analytic review. *Psychonomic Bulletin & Review* 4 (2): 184–197.

Boltz, M. G. 1998. The processing of temporal and nontemporal information in the remembering of event durations and musical structure. *Journal of Experimental Psychology: Human Perception and Performance* 24 (4): 1087–1104.

Bonner, S. E. 1994. A model of the effects of audit task complexity. *Accounting, Organizations and Society* 19 (3): 213–234.

———, R. Hastie, G. B. Sprinkle, and S. M. Young. 2000. A review of the effects of financial incentives on performance in laboratory tasks: Implications for management accounting. *Journal of Management Accounting Research* 12: 19–64.

———, and G. B. Sprinkle. 2002. The effects of monetary incentives on effort and task performance: Theories, evidence and a framework for research. *Accounting, Organizations and Society* 27 (4–5): 303–345.

Brown, S. W. 1985. Time perception and attention: The effects of prospective versus retrospective paradigms and task demands on perceived duration. *Perception & Psychophysics* 38 (2): 115–124.

———, D. A. Stubbs, and A. N. West. 1992. Attention, multiple timing, and psychophysical scaling of temporal judgments. In *Time, Action and Cognition: Towards Bridging the Gap*, edited by F. Macar, W. J. Friedman, and V. Pouthas. New York, NY: Springer.

———. 1997. Attentional resources in timing: Interference effects in concurrent temporal and nontemporal working memory tasks. *Perception & Psychophysics* 59 (7): 1118–1140.

———, and M. G. Boltz. 2002. Attentional processes in time perception: Effects of mental workload and event structure. *Journal of Experimental Psychology* 28 (3): 600–615.

Buckless, F. A., and S. P. Ravenscroft. 1990. Contrast coding: A refinement of ANOVA in behavioral analysis. *The Accounting Review* 65 (4): 933–945.

Callahan, C. M., and A. E. Gabriel. 1998. The differential impact of accurate product cost information in imperfectly competitive markets: A theoretical and empirical investigation. *Contemporary Accounting Research* 15 (4): 419–455.

Christensen, J., and J. Demski. 1995. The classical foundations of modern costing. *Management Accounting Research* 6: 13–32.

Cooper, R., and R. S. Kaplan. 1988. Measure costs right: Make the right decisions. *Harvard Business Review* 66 (5): 96–103.

———, ———, L. S. Maisel, E. Morrissey, and R. M. Oehm. 1992. *Implementing Activity-Based Cost Management: Moving From Analysis to Action.* Montvale, NJ: Institute of Management Accountants.

———, and ———. 1998. *Design of Cost Management Systems.* 2nd edition. Englewood Cliffs, NJ: Prentice Hall.

Datar, S. M., and M. Gupta. 1994. Aggregation, specification and measurement errors in product costing. *The Accounting Review* 69 (4): 567–591.

Drewnowski, A. 1980. Attributes and priorities in short-term recall: A new model of memory span. *Journal of Experimental Psychology: General* 109: 208–250.

Ellis-Newman, J., and P. Robinson. 1998. The cost of library services: Activity-based costing in an Australian academic library. *The Journal of Academic Librarianship* 24 (5): 373–379.

Fiedler, K., and T. Armbruster. 1994. Two halves may be more than one whole: Category-split effects on frequency illusions. *Journal of Personality and Social Psychology* 66 (4): 633–645.

Flückiger, M. 1985. The effects of time pressure on duration discrimination. In *Time, Mind, and Behavior*, edited by J. E. Michon, and J. L. Jackson. New York, NY: Springer.

Francis-Smythe, J. A., and I. T. Robertson. 1999. On the relationship between time management and time estimation. *British Journal of Psychology* 90: 333–347.

Goebel, D. J., G. W. Marshall, and W. B. Locander. 1998. Activity-based costing: Accounting for a market orientation. *Industrial Marketing Management* 27 (6): 497–510.

Gorn, G. J., A. Chattopadhyay, J. Sengupta, and S. Tripathi. 2004. Waiting for the web: How screen color affects time perception. *Journal of Marketing Research* 41 (21): 215–225.

Grunfeld, Y., and Z. Griliches. 1960. Is aggregation necessarily bad? *The Review of Economics and Statistics* 42 (1): 1–13.

Gupta, M., and R. R. King. 1997. An experimental investigation of the effect of cost information and feedback on product cost decisions. *Contemporary Accounting Research* 14: 99–127.

Hasher, L., and R. T. Zacks. 1979. Automatic and effortful processes in memory. *Journal of Experimental Psychology: General* 108 (3): 356–388.

Ittner, C. D. 1999. Activity-based costing concepts for quality improvement. *European Management Journal* 17 (5): 492–500.

Jones, M. R., M. G. Boltz, and J. M. Klein. 1993. Expected endings and judged duration. *Memory & Cognition* 21 (5): 646–665.

Kahneman, D. 1973. *Attention and Effort.* Englewood Cliffs, NJ: Prentice Hall.

Kaplan, R. S., and A. A. Atkinson. 1998. *Advanced Management Accounting.* Englewood Cliffs, NJ: Prentice Hall.

———, and S. R. Anderson. 2004. Time-driven activity-based costing. *Harvard Business Review* (November): 131–138.

Kobeissi, B. J., M. Gupta, C. A. Perez, N. Dopuch, J. M. Michalski, and G. Van Antwerp. 1998. Physician resource utilization in radiation oncology: A model based on management of carcinoma of the prostate. *International Journal of Radiation Oncology & Biology & Physics* 40 (3): 593–603.

Libby, R., and M. G. Lipe. 1992. Incentives, effort, and the cognitive processes involved in accounting-related judgments. *Journal of Accounting Research* 30 (2): 249–273.

Lim, S. S., and S. Sunder. 1990. Accuracy of linear valuation rules in industry-segmented environments. *Journal of Accounting and Economics* 13: 167–188.

———, and ———. 1991. Efficiency of asset valuation rules under price movement and measurement errors. *The Accounting Review* 66 (4): 669–693.

Maher, M. W., and M. L. Marais. 1998. A field study on the limitations of activity-based costing when resources are provided on a joint and indivisible basis. *Journal of Accounting Research* 36 (1): 129–142.

Major, M., and T. Hopper. 2005. Implementing ABC in a Portuguese telecommunications company. *Management Accounting Research* 16 (2): 205–229.

Merchant, K., and M. Shields. 1993. When and why to measure costs less accurately to improve decision making. *Accounting Horizons* 7 (2): 76–81.

Migliore, L., L. Messineo, and M. Cardaci. 2000. A model of the effects of cognitive load on the subjective estimation and production of time intervals. *Biosystems* 58: 187–193.

Miller, G. A. 1956. The magical number seven plus or minus two: Some limits on our capacity for processing information. *Psychological Review* 63: 81–97.

Narayanan, V. G., and R. G. Sarkar. 2002. The impact of activity-based costing on managerial decisions at Insteel Industries—A field study. *Journal of Economics & Management Strategy* 11 (2): 257–288.

Noreen, E. 1991. Conditions under which activity-based cost systems provide relevant costs. *Journal of Management Accounting Research* 3 (Fall): 159–169.

Ornstein, R. E. 1969. *On the Experience of Time.* Baltimore, MD: Penguin Books.

Poynter, W. D. 1983. Duration judgments and the segmentation of experience. *Memory & Cognition* 11: 77–82.

Rose, J. M., and C. J. Wolfe. 2000. The effects of system design alternatives on the acquisition of tax knowledge from a computerized tax decision aid. *Accounting, Organizations and Society* 25: 285–306.

Vera-Muñoz, S. C. 1998. The effects of accounting knowledge and context on the omission of opportunity costs in resource allocation decisions. *The Accounting Review* 73 (1): 47–72.

Waller, W., B. Shapiro, and G. Sevcik. 1999. Do cost-based pricing biases persist in laboratory markets? *Accounting, Organizations and Society* 24 (8): 717–739.

Zakay, D. 1993. The impact of time perception processes on decision making under time-stress. In *Time Pressure and Stress in Human Judgment and Decision Making*, edited by O. Svenson, and A. J. Maule. New York, NY: Plenum Press.

———, and T. Feldman. 1993. The role of segmentation and recallability in retrospective time estimation. *Psychological Record* 43 (3): 415–428.

———, and J. Shub. 1998. Concurrent duration production as a workload measure. *Ergonomics* 41 (8): 1115–1128.

THE ACCOUNTING REVIEW
Vol. 83, No. 3
2008
pp. 757–787

Real and Accrual-Based Earnings Management in the Pre- and Post-Sarbanes-Oxley Periods

Daniel A. Cohen
New York University

Aiyesha Dey
University of Chicago

Thomas Z. Lys
Northwestern University

ABSTRACT: We document that accrual-based earnings management increased steadily from 1987 until the passage of the Sarbanes-Oxley Act (SOX) in 2002, followed by a significant decline after the passage of SOX. Conversely, the level of real earnings management activities declined prior to SOX and increased significantly after the passage of SOX, suggesting that firms switched from accrual-based to real earnings management methods after the passage of SOX. We also document that the accrual-based earnings management activities were particularly high in the period immediately preceding SOX. Consistent with these results, we find that firms that just achieved important earnings benchmarks used less accruals and more real earnings management after SOX when compared to similar firms before SOX. In addition, our analysis provides evidence that the increases in accrual-based earnings management in the period preceding SOX were concurrent with increases in equity-based compensation. Our results suggest that stock-option components provide a differential set of incentives with regard to accrual-based earnings management. We document that while new options granted during the current period are negatively associated with income-increasing accrual-based earnings management, unexercised options are positively associated with income-increasing accrual-based earnings management.

We acknowledge the financial support from the Accounting Research Center at the Kellogg School and helpful comments from Dan Dhaliwal, April Klein, Krishna Kumar, Eddie Riedl, Suraj Srinivasan, Ira Weiss, Jerry Zimmermann, two anonymous reviewers, participants at the HBS Accounting and Control seminar, the 2006 FARS Meeting, the 2004 AAA Annual Meeting, the 2004 Corporate Governance Conference at the University of Washington in St. Louis, and the 2004 LBS Accounting Symposium. All remaining errors are our own.

Editor's note: This paper was accepted by Dan Dhaliwal.

Submitted February 2006
Accepted October 2007

I. INTRODUCTION

The recent wave of corporate governance failures has raised concerns about the integrity of the accounting information provided to investors and resulted in a drop in investor confidence (Jain et al. 2003; Jain and Rezaee 2006; Rezaee 2004). These failures were highly publicized and ultimately led to the passage of the Sarbanes-Oxley Act (SOX, July 30, 2002). The changes mandated by SOX were extensive, with President George W. Bush commenting that this Act constitutes "the most far-reaching reforms of American business practices since the time of Franklin D. Roosevelt."[1] Similarly, the head of the AICPA commented that SOX "contains some of the most far-reaching changes that Congress has ever introduced to the business world"[2] including an unprecedented shift in the regulation of corporate governance from the states to the federal government.[3]

Although SOX proposed sweeping changes, the scope of the events that led to the passage of the Act and the consequences of the resulting regulatory changes have yet to be systematically studied. Specifically, it is unclear whether there really was a widespread breakdown of the reliability of financial reporting prior to the passage of SOX or whether the highly publicized scandals were isolated instances of individuals engaging in blatant financial manipulations. If it were the former, then how did the passage of SOX affect firms' financial reporting practices? Moreover, some argue that these frauds occurred after 70 years of ever-increasing securities regulation, suggesting that more regulation may not be the answer (Ribstein 2002).

We investigate the prevalence of both accrual-based and real earnings management activities in the period leading to the passage of SOX and in the period following the passage of SOX. Our primary motivation for conducting this analysis is to investigate whether the period leading to the passage of SOX was characterized by a widespread increase in earnings management rather than by a few highly publicized events, and whether the passage of SOX resulted in a reduction in earnings management.

We carry out our investigation by dividing the sample period into two time periods: the period prior to the passage of SOX (the pre-SOX period: 1987 through 2001), and the period after the passage of SOX (the post-SOX period: 2002 through 2005). We further subdivide the pre-SOX period into two subperiods: the period prior to the major corporate scandals (the pre-SCA period: 1987 through 1999) and the period immediately preceding the passage of SOX when the major scandals occurred (the SCA period: 2000 and 2001).

We document that the pre-SOX period was characterized by increasing accrual-based earnings management, culminating in even larger increases in the SCA period, but declining real earnings management. We also document that the increase in accrual-based earnings management in the SCA period was associated with a contemporaneous increase in option-based compensation. However, option-based compensation exhibits two opposing incentives with respect to current stock prices (and thus earnings management). On the one hand, managers have an incentive to lower the current stock price around stock option award (e.g., Yermack 1997; Aboody and Kasznik 2000). On the other hand, for unexercised options (excluding new option grants), managers have an incentive to increase current stock prices and, hence, to manage earnings upward. To capture these two effects, we partition overall options held at the end of the year into unexercised options (excluding new options grants), new options granted in the current period, and exercisable vested options. Consistent with the notion that these components provide a differential set of incentives, we find that while new options granted during the current period are negatively associated with income-increasing discretionary accruals, unexercised options are positively associated with income-increasing discretionary accruals. However, consistent with Cheng and Warfield (2005), we do not find a significant association between accrual-based earnings management and exercisable vested options. Furthermore, we document that the relation between income-increasing discretionary accruals and unexercised options (excluding new option grants) declined in the post-SOX period. To the best of our knowledge, we are the first to document this evidence empirically.

---

[1] Bumiller (2002).

[2] Melancon (2002).

[3] Traditionally, the federal government has focused on regulating disclosure, public trading, and antitrust, while regulating corporate governance has been the focus of the states.

Following the passage of SOX accrual-based earnings management declined significantly, while real earnings management increased significantly. At the same time, option-based compensation decreased. Consistent with the results of a recent survey by Graham et al. (2005), this suggests that firms switched to managing earnings using real methods, possibly because these techniques, while more costly, are likely to be harder to detect.

In additional analyses, we examine both the pre-SOX and post-SOX accrual-based and real earnings management activities for a subset of firms that are more likely to have managed earnings (we refer to these firms as the SUSPECT firms). Specifically, we examine three incentives for managing earnings, namely, meeting or beating last year's earnings, meeting or beating the consensus analysts' forecast and avoiding reporting losses. We focus on these incentives because Graham et al. (2005) document that these specific motives are among the most important reasons for earnings management behavior.

Consistent with our full sample results, we find that both before and after SOX, SUSPECT firms had significantly higher discretionary accruals when compared to firms that either just missed those benchmarks or firms that were not close to either making or missing those benchmarks. However, these SUSPECT firms used significantly less income-increasing accrual-management after SOX when compared to firms in similar circumstances before SOX. An analysis of the real earnings management behavior of these firms indicates that the SUSPECT firms had significantly higher real earnings management activities after SOX when compared to the firms in similar circumstances prior to SOX.

Our results contribute to the current debate on the pervasiveness of earnings management prior to the passage of SOX, and the impact of SOX on such behavior. While we find an increase in accrual-based earnings management prior to SOX, our evidence suggests that firms are likely to have switched to earnings-management techniques that, while likely to be more costly to shareholders, are harder to detect. This evidence forms an important consideration in the debate on the costs and benefits of the new regulation.

The remainder of the paper proceeds as follows. Section II provides a discussion of the research questions and hypotheses. Section III discusses the empirical methodology, including the data and sample selection and the various measures of earnings management used in the study. The tests and results are discussed in Section IV, and Section V concludes.

II. MOTIVATION, RESEARCH QUESTIONS, AND HYPOTHESES

In a recent commentary, U.S. Treasury Secretary Henry Paulson emphasized the importance of strong capital markets, and pointed out that capital markets rely on trust and that trust is based on financial information presumed to be accurate and to reflect economic reality.[4] The series of corporate scandals occurring in 2000–2001 eroded that trust in financial reports. Indeed, one of the main objectives of SOX was to restore the integrity of financial statements by curbing earnings management and accounting fraud. Therefore, the extent of earnings management prior to SOX and the effect of SOX on earnings management is an important research topic.

---

[4] Paulson (2007).

The primary purpose of this paper is to examine the extent of earnings management in the period leading to the scandals and prior to SOX, and the changes in such activities after the passage of SOX.[5] Our examination of changes in firms' earnings management activities is motivated in part by the literature documenting that managerial propensity to manage earnings and to avoid negative earnings surprises has increased significantly over time (e.g., Brown 2001; Bartov et. al. 2002; Lopez and Rees 2001; Matsumoto 2002; Brown and Caylor 2005). Our main objective is to examine whether the degree of earnings management increased over time and reached a zenith in the period surrounding the corporate accounting scandals, and declined after the passage of SOX.

Consistent with the literature, we examine earnings management activities using discretionary accruals. However, in addition to using accrual-based accounting estimates and methods, firms are likely to employ real operational activities to manipulate earnings numbers as well (e.g., Healy and Wahlen 1999; Dechow and Skinner 2000). In fact, in their survey Graham et al. (2005) report the following:[6]

> [W]e find strong evidence that managers take real economic actions to maintain accounting appearances. In particular, 80% of survey participants report that they would decrease discretionary spending on R&D, advertising, and maintenance to meet an earnings target. More than half (55.3%) state that they would delay starting a new project to meet an earnings target, even if such a delay entailed a small sacrifice in value.

Thus, to provide a more complete study of the trends in earnings management activities in the periods before and after SOX, we also examine real earnings management activities over the sample period.

Next, we examine possible explanations for any changes in earnings management activities over the sample period. We focus on the hypothesis that managers' choices of accounting practices are influenced by the impact of these accounting methods on their compensation. Managers with higher stock- and option-based compensation are more sensitive to short-term stock prices, and can use their discretion to affect reported earnings if capital markets have difficulty detecting earnings management (see Fields et al. [2001] for a discussion on the pricing of earnings management.)[7]

---

[5] In a related study, Lobo and Zhou (2006) investigate whether the SOX mandate that financial statements be certified by firms' CEOs and CFOs resulted in an increase in the conservatism in financial reporting, and find evidence of greater conservatism in reported earnings.

[6] In a recent working paper, Zang (2006) investigates whether managers use real and accrual manipulations in managing earnings as substitutes. Based on a model she develops, she provides evidence consistent with managers using real and accrual manipulations as substitutes.

[7] For instance, Coffee (2003) asserts that the increase in stock-based executive compensation created an environment where managers became very sensitive to short-term stock performance. Greenspan (2002) opines that "the highly desirable spread of shareholding and options among business managers perversely created incentives to artificially inflate earnings to keep stock prices high and rising." Fuller and Jensen (2002, 42) also state that "[a]s stock options became an increasing part of executive compensation, and managers who made great fortunes on options became the stuff of legends, the preservation or enhancement of short-term stock prices became a personal (and damaging) priority for many CEOs and CFOs. High share prices and earnings multiples stoked already amply endowed managerial egos, and management teams proved reluctant to undermine their own stature by surrendering hard won records of quarter-over-quarter earnings growth."

Prior studies (e.g., Cheng and Warfield 2005; Bergstresser and Philippon 2006) provide evidence suggesting that equity incentives derived from stock-option compensation are positively associated with managements' likelihood to engage in accrual-based earnings management activities. However, Johnson et al. (2005) conclude that only unrestricted stock holdings are associated with the occurrence of accounting fraud; the stock option grants are not. Further, Erickson et al. (2006) find no consistent evidence that executive equity incentives are associated with fraud.

In addition to equity-based compensation, executives are also rewarded based on explicit bonus-linked targets for reported income. Healy (1985) presents evidence that the accruals policies of managers are related to the nonlinear incentives inherent in their bonus contracts. Therefore, we investigate whether earnings-based compensation contracts are associated with earnings management.

Our focus on the compensation structure is motivated by the current debate whether option-based compensation and bonus grants are associated with earnings management. Further, there has been a significant increase in the grant of stock options in the past decade. We examine whether those increases (prior to SOX) and decreases (after SOX) are related to the level of earnings management during that period.

Specifically, our hypothesis that managers behave opportunistically due to compensation-related incentives has two empirical predictions. First, changes in reported earnings are affected by changes in the compensation and incentives of managers. Second, even after controlling for managerial incentives, earnings management would decline after the passage of SOX, either because of the sanctions imposed on managers by SOX or because of the adverse publicity and legal costs imposed on executives and firms who were accused of questionable or even fraudulent reporting practices.

Our final objective is to investigate whether after the passage of SOX corporations replaced some accrual-based earnings management with real earnings management, which is not only likely to be harder to detect, but also likely to be more costly to the firm (Graham et al. 2005).[8] We based this expectation on the premise that after the passage of SOX accrual manipulations were more likely to draw auditors' or regulators' scrutiny than real earnings management. If firms were more wary after the passage of SOX, then they would be more likely to substitute real earnings management for accrual-based earnings management after SOX. This conjecture is also suggested by Graham et al. (2005):

> [W]e acknowledge that the aftermath of accounting scandals at Enron and WorldCom and the certification requirements imposed by the Sarbanes-Oxley Act may have changed managers' preferences for the mix between taking accounting versus real actions to manage earnings.

Given the above argument, investigating the trends in both real and accrual-based earnings management after SOX is important. Evidence of a decline in one type of earnings management may lead one to conclude that such activities have decreased in response to regulators or other events, when in fact a substitution of one earnings management method for another has occurred. We thus hypothesize and test whether the level of real earnings management increased after the passage of SOX, i.e., whether firms substituted between real and accrual-based methods after SOX. The next section discusses the empirical methodology employed in the study.

---

[8] In their survey, Graham et al. (2005, 66) report: "Managers candidly admit that they would take real economic actions such as delaying maintenance or advertising expenditure, and would even give up positive NPV projects, to meet short-term earnings benchmarks."

III. EMPIRICAL METHODOLOGY

Data and Sample Description

We collect our sample from the Compustat annual industrial and research files for the period 1987–2005. We restrict our sample to all nonfinancial firms with available data, and require at least eight observations in each two-digit SIC grouping per year. Further, we require that each firm-year observation has the data necessary to calculate the discretionary accruals metrics and real earnings management proxies we employ in our analysis. This restriction likely introduces a survivorship bias, biasing the sample toward larger and more successful firms. However, we expect that this bias will reduce the variation in our earnings management metric, resulting in a more conservative test of our research questions.

Following Collins and Hribar (2002), we use cash flows from operations obtained from the Statement of Cash Flows reported under the Statement of Financial Accounting Standards No. 95 (SFAS No. 95, FASB 1987).[9] The sample period of 1987–2005 permits us to use SFAS No. 95 statement of cash flow data to estimate accruals, rather than a balance sheet approach.

The sample obtained from Compustat consists of 8,157 firms representing 87,217 firm-year observations. To test the compensation hypothesis, we use data from ExecuComp, which is available only from 1992 onward. Thus, merging the full sample with ExecuComp results in a second (and smaller) sample consisting of 2,018 firms and 31,668 firm-year observations (the ExecuComp sample) for the 1992 through 2005 period.

Event Periods

We focus in our analysis on earnings management across two main time periods—the pre-SOX period (further classified into the pre-SCA and the SCA periods), and the post-SOX period. The pre-SOX period extends from 1987 through 2001, and the post-SOX period extends from 2002 through the end of 2005. Within the pre-SOX period, we classify the period from 1987 through 1999 as the pre-SCA period, and the period from 2000 through 2001 as the SCA period (i.e., the period that purportedly lead to the passage of SOX).[10] Figure 1 depicts these different time periods analyzed.[11]

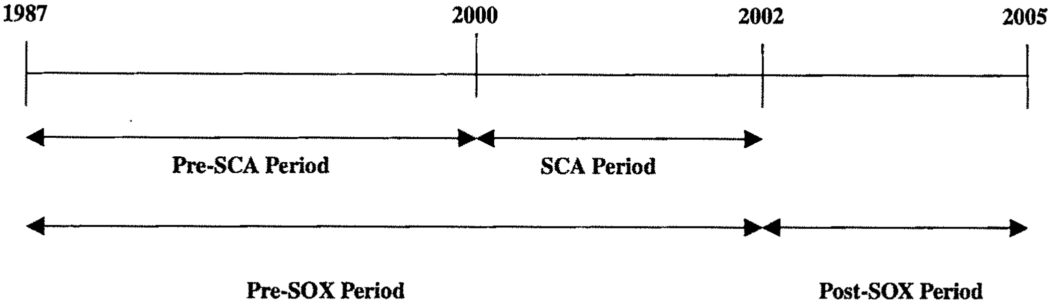
---

[9] SFAS No. 95 requires firms to present a statement of cash flows for fiscal years ending after July 15, 1988. Some firms early-adopted SFAS No. 95, so our sample begins in 1987.

[10] We acknowledge that the subdivision into the pre-SCA and SCA periods may induce hindsight bias into the analysis. We thus repeat our analysis by only dividing the entire sample period into the pre-SOX and the post-SOX periods. Our conclusions on earnings management activities before and after SOX are unchanged and become stronger. We do not report these results in the paper for the sake of brevity, but will provide them upon request.

[11] The use of annual data determines how we subdivide the sample period to some extent. Although some scandals took place in the beginning of 2002, we include 2002 in the post-SOX period, since SOX was passed in 2002. Also, even though the most public phase of the scandals began with Enron in 2001, we include year 2000 in the SCA period since some frauds also occurred in 2000 (e.g., Xerox). Further, as a robustness check, we also repeat our analysis using quarterly data and subdivide the sample period using the "Corporate Scandal Sheet" developed by Forbes (2002). Specifically, we define the SCA period as extending from Q3, 2001 through Q2, 2002, and the post-SOX period as extending from Q3, 2002 onward. Our main conclusions remain unchanged, which provides added confidence in the results obtained using annual data. Finally, as suggested by the referee we repeat all analyses by defining the post-SOX period as the years 2003 through 2005 and this did not materially alter any of our reported results.

FIGURE 1
Time Periods Analyzed



Earnings Management Metrics

Accrual-Based Earnings Management

We use a cross-sectional model of discretionary accruals, where for each year we estimate the model for every industry classified by its two-digit SIC code. Thus, our approach partially controls for industry-wide changes in economic conditions that affect total accruals while allowing the coefficients to vary across time (e.g., Kasznik 1999; DeFond and Jiambalvo 1994).[12]

Our primary model is the modified cross-sectional Jones model (Jones 1991) as described in Dechow et al. (1995).[13] The modified Jones model is estimated for each two-digit SIC-year grouping as follows:

$$\frac{TA_{it}}{Assets_{i,t-1}} = k_{1t}\frac{1}{Assets_{i,t-1}} + k_2\frac{\Delta REV_{it}}{Assets_{i,t-1}} + k_3\frac{PPE_{it}}{Assets_{i,t-1}} + \varepsilon_{it} \quad (1)$$

where, for fiscal year t and firm i, TA represents total accruals defined as:

TA_{it} = $EBXI_{it} - CFO_{it}$, where $EBXI$ is the earnings before extraordinary items and discontinued operations (annual Compustat data item 123) and CFO is the operating cash flows (from continuing operations) taken from the statement of cash flows (annual Compustat data item 308 − annual Compustat data item 124);
$Assets_{it-1}$ = total assets (annual Compustat data item 6);
ΔREV_{it} = change in revenues (annual Compustat data item 12) from the preceding year; and

---

[12] We obtain qualitatively the same results when we use a time-series approach that assumes temporal stationarity of the parameters for each firm.

[13] Caveat: Various studies in the literature raise the concern that discretionary accruals measured using the Jones model might be capturing nondiscretionary components and these errors in discretionary accruals are likely to be correlated with stock prices and performance measures in general. While this concern is valid and we acknowledge this limitation in measuring discretionary accruals, note that we use discretionary accruals as a dependent variable and not as an explanatory variable. If indeed discretionary accruals are measured with error, then the only consequence in our case will be a lower explanatory power of the model, i.e., we will obtain lower R^2s. Otherwise, using discretionary accruals measured using the Jones model as a dependent variable is not likely to introduce any bias in our results.

PPE_{it} = gross value of property, plant, and equipment (annual Compustat data item 7).

The coefficient estimates from Equation (1) are used to estimate the firm-specific normal accruals (NA_{it}) for our sample firms:

$$NA_{it} = \hat{k}_{1t}\frac{1}{Assets_{i,t-1}} + \hat{k}_2\frac{(\Delta REV_{it} - \Delta AR_{it})}{Assets_{i,t-1}} + \hat{k}_3\frac{PPE_{it}}{Assets_{i,t-1}} \tag{2}$$

where ΔAR_{it} is the change in accounts receivable (annual Compustat data item 2) from the preceding year. Following the methodology used in the literature, we estimate the industry-specific regressions using the change in reported revenues, implicitly assuming no discretionary choices with respect to revenue recognition. However, while computing the normal accruals, we adjust the reported revenues of the sample firms for the change in accounts receivable to capture any potential accounting discretion arising from credit sales. Our measure of discretionary accruals is the difference between total accruals and the fitted normal accruals, defined as $DA_{it} = (TA_{it}/Assets_{it-1}) - NA_{it}$.

In contrast to studies that focus on a specific corporate event, our analysis using the full sample is conducted in calendar time. Consequently, because accruals reverse over time and we cannot condition the analysis on events that are hypothesized to provide managers with incentives to manage reported earnings in any given direction (e.g., inflate reported earnings) we compute the absolute value of discretionary accruals to proxy for earnings management and refer to it as *ABS_DA* throughout the analysis.[14] In contrast, our test of the SUSPECT firms (that is, firms that were just able to meet or beat earnings benchmarks) is based on a directional test.

In our robustness tests, we used two alternative measures of discretionary accruals. In one alternative measure we estimated the following in the first stage:

$$\frac{TA_{it}}{Assets_{i,t-1}} = k_{1t}\frac{1}{Assets_{i,t-1}} + k_2\frac{(\Delta REV_{it} - \Delta AR_{it})}{Assets_{i,t-1}} + k_3\frac{PPE_{it}}{Assets_{i,t-1}} + \varepsilon_{it}. \tag{3}$$

Using the coefficient estimates obtained from Equation (3), we calculated the level of normal accruals (NA_{it}) as a percent of lagged total assets. We also repeat our tests by using a measure based on the performance-matched discretionary accruals advanced in Kothari et al. (2005). As suggested by Kothari et al. (2005), we match each firm-year observation with another from the same two-digit SIC code and year with the closest return on assets in the current year, ROA_{it} (net income divided by total assets).[15] Our results using these alternate measures of accruals are consistent with those reported in the paper.

Real Earnings Management

We rely on prior studies to develop our proxies for real earnings management. As in Roychowdhury (2006) we consider the abnormal levels of cash flow from operations (*CFO*), discretionary expenses and production costs to study the level of real activities manipulations. Subsequent studies, such as Zang (2006) and Gunny (2005), provide evidence of the construct validity of these proxies. We focus on three manipulation methods and their impact on the above three variables:

---

[14] We repeat our analysis using the square of discretionary accruals and the results are somewhat stronger. Although the squared discretionary accruals have more desirable distributional properties, we report the results using the absolute values of discretionary accruals to allow comparison with other research.

[15] We also carry out performance matching based on two-digit SIC code, year, and ROA (both current ROA and lagged ROA) and obtain results similar to those reported in the paper.

1. Acceleration of the timing of sales through increased price discounts or more lenient credit terms. Such discounts and lenient credit terms will temporarily increase sales volumes, but these are likely to disappear once the firm reverts to old prices. The additional sales will boost current period earnings, assuming the margins are positive. However, both price discounts and more lenient credit terms will result in lower cash flows in the current period.
2. Reporting of lower cost of goods sold through increased production. Managers can increase production more than necessary in order to increase earnings. When managers produce more units, they can spread the fixed overhead costs over a larger number of units, thus lowering fixed costs per unit. As long as the reduction in fixed costs per unit is not offset by any increase in marginal cost per unit, total cost per unit declines. This decreases reported cost of goods sold (*COGS*) and the firm can report higher operating margins. However, the firm will still incur other production and holding costs that will lead to higher annual production costs relative to sales, and lower cash flows from operations given sales levels.
3. Decreases in discretionary expenses that include advertising expense, research and development, and SG&A expenses. Reducing such expenses will boost current period earnings. It could also lead to higher current period cash flows (at the risk of lower future cash flows) if the firm generally paid for such expenses in cash.

We first generate the normal levels of *CFO*, discretionary expenses, and production costs using the model developed by Dechow et al. (1998) as implemented in Roychowdhury (2006). We express normal *CFO* as a linear function of sales and change in sales. To estimate this model, we run the following cross-sectional regression for each industry and year:

$$\frac{CFO_{it}}{Assets_{i,t-1}} = k_{1t}\frac{1}{Assets_{i,t-1}} + k_2\frac{Sales_{it}}{Assets_{i,t-1}} + k_3\frac{\Delta Sales_{it}}{Assets_{i,t-1}} + \varepsilon_{it}. \tag{4}$$

Abnormal *CFO* is actual *CFO* minus the normal level of *CFO* calculated using the estimated coefficient from Equation (4).

Production costs are defined as the sum of *COGS* and change in inventory during the year. We model *COGS* as a linear function of contemporaneous sales:

$$\frac{COGS_{it}}{Assets_{i,t-1}} = k_{1t}\frac{1}{Assets_{i,t-1}} + k_2\frac{Sales_{it}}{Assets_{i,t-1}} + \varepsilon_{it}. \tag{5}$$

Next, we model inventory growth by the following:

$$\frac{\Delta INV_{it}}{Assets_{i,t-1}} = k_{1t}\frac{1}{Assets_{i,t-1}} + k_2\frac{\Delta Sales_{it}}{Assets_{i,t-1}} + k_3\frac{\Delta Sales_{i,t-1}}{Assets_{i,t-1}} + \varepsilon_{it}. \tag{6}$$

Using Equations (5) and (6), we estimate the normal level of production costs as:

$$\frac{Prod_{it}}{Assets_{i,t-1}} = k_{1t}\frac{1}{Assets_{i,t-1}} + k_2\frac{Sales_{it}}{Assets_{i,t-1}} + k_3\frac{\Delta Sales_{it}}{Assets_{i,t-1}} + k_4\frac{\Delta Sales_{it-1}}{Assets_{i,t-1}} + \varepsilon_{it}. \quad (7)$$

We model the normal level of discretionary expenses as:

$$\frac{DiscExp_{it}}{Assets_{i,t-1}} = k_{1t}\frac{1}{Assets_{i,t-1}} + k_2\frac{Sales_{it}}{Assets_{i,t-1}} + \varepsilon_{it}. \quad (8)$$

Modeling discretionary expenses as a function of current sales creates a mechanical problem if firms manage sales upward to increase reported earnings in a certain year, resulting in significantly lower residuals from running a regression as derived in Equation (8). To address this issue, we model discretionary expenses as a function of lagged sales and estimate the following model to derive 'normal' levels of discretionary expenses:

$$\frac{DiscExp_{it}}{Assets_{i,t-1}} = k_{1t}\frac{1}{Assets_{i,t-1}} + k_2\frac{Sales_{i,t-1}}{Assets_{i,t-1}} + \varepsilon_{it}. \quad (9)$$

In the above equations *CFO* is cash flow from operations in period *t* (Compustat data item 308 − annual Compustat data item 124); *Prod* represents the production costs in period *t*, defined as the sum of *COGS* (annual Compustat data item 41) and the change in inventories (annual Compustat data item 3); and *DiscExp* represents the discretionary expenditures in period *t*, defined as the sum of advertising expenses (annual Compustat data item 45), R&D expenses (annual Compustat data item 46),[16] and SG&A (annual Compustat data item 189). The abnormal CFO (*R_CFO*), abnormal production costs (*R_PROD*) and abnormal discretionary expenses (*R_DISX*) are computed as the difference between the actual values and the normal levels predicted from Equations (4), (7), and (9). We use these three variables as proxies for real earnings management.

Given sales levels, firms that manage earnings upward are likely to have one or all of these: unusually low cash flow from operations, and/or unusually low discretionary expenses, and/or unusually high production costs. In order to capture the effects of real earnings management through all these three variables in a comprehensive measure, we compute a single variable by combining the three individual real earnings management variables. Specifically, we compute *RM_PROXY* as the sum of the standardized variables, *R_CFO*, *R_PROD*, and *R_DISX*. However, we acknowledge that the three individual variables have different implications for earnings that may dilute any results using *RM_PROXY* alone. We thus report results corresponding to the single real earnings management proxy (*RM_PROXY*) as well as the three individual real earnings management proxies (*R_CFO*, *R_PROD*, and *R_DISX*).

IV. TESTS AND RESULTS

Descriptive Statistics: Earnings Management

We begin with an exploratory analysis of the trends over time in the various earnings management metrics. Table 1, Panel A provides summary statistics of the full sample, while

---

[16] As long as SG&A is available, advertising expenses and R&D are set to zero if they are missing.

TABLE 1
Descriptive Statistics

	25th Percentile	Mean	Median	75th Percentile	Standard Deviation
Panel A: Full Sample, 1987–2005 (n = 87,217)					
Total Assets	17.54	1401.43	89.02	438.68	8815
Market Capitalization	15.08	1616.11	76.71	441.13	11855.55
Sales	17.05	1396.32	94.70	492.98	7555.65
Growth of Sales	−0.04	0.13	0.08	0.25	0.49
OC (Days)	79.15	142.56	124.05	186.01	90.89
Leverage	0.04	0.41	0.21	0.39	0.26
Market-to-Book	1.11	4.94	1.92	3.48	6.17
Total Accruals	−0.12	−0.10	−0.06	−0.001	0.25
DA	−0.05	0.00	0.00	0.05	0.20
Positive_DA	0.02	0.09	0.06	0.11	0.09
Negative_DA	−0.03	−0.05	−0.02	−0.01	0.18
ABS_DA	0.03	0.11	0.06	0.12	0.17
R_CFO	−0.06	−0.02	0.01	0.08	0.35
R_PROD	−0.17	−0.06	0.04	0.06	0.31
R_DISX	−0.09	0.08	0.01	0.18	0.58
RM_PROXY	−0.08	0.01	−0.01	0.06	0.35
Panel B: ExecuComp Sample, 1992–2005 (n = 31,668)					
Total Assets	334.88	4232.66	849.40	2611.80	19038.10
Market Capitalization	388.27	5491.55	978.67	3280.46	20283.20
Sales	331.89	4025.76	895.94	2678.90	13039.01
Growth of Sales	0.01	0.18	0.09	0.22	0.76
OC (Days)	72.49	124.69	111.27	160.10	73.61
Leverage	0.05	0.22	0.20	0.33	0.20
Market-to-Book	1.63	4.70	2.48	4.04	6.45
Total Accruals	−0.10	−0.07	−0.06	−0.02	0.12
DA	−0.05	−0.01	−0.01	0.03	0.11
Positive_DA	0.00	0.03	0.00	0.03	0.06
Negative_DA	−0.05	−0.04	−0.01	−0.00	0.08
ABS_DA	0.01	0.07	0.04	0.08	0.09
R_CFO	−0.00	0.05	0.05	0.11	0.15
R_PROD	−0.17	−0.05	−0.04	0.05	0.28
R_DISX	−0.10	0.01	0.00	0.12	0.15
RM_PROXY	−0.06	0.00	0.00	0.06	0.16
BONUS (%)	0.04	0.16	0.14	0.25	0.15
BONUS ($)	60.00	454.04	219.32	526.80	821.30
EX_OPTIONS (%)	0.312	0.745	0.432	0.874	0.857
UN_OPTIONS (%)	0.072	0.212	0.102	0.347	0.431
GRNT_OPTION (%)	0.032	0.129	0.049	0.128	0.304
OWNER (%)	0.764	5.191	1.067	3.057	1.067

(*continued on next page*)

TABLE 1 (continued)

Variable Definitions:

Total Assets = annual Compustat data item 6;

Market Capitalization = the price per share (annual Compustat data item 199) times the number of shares outstanding (annual Compustat data item 25);

Sales = annual Compustat data item 12;

Growth of Sales = the change in sales divided by lagged sales;

Total Accruals = the difference between operating cash flows (annual Compustat data item 308), adjusted for extraordinary items and discontinued operations (annual Compustat data item 124) and income before extraordinary items (annual Compustat data item 123) divided by lagged total assets;

DA = discretionary accruals computed using the Modified Jones Model;

ABS_DA = the absolute value of discretionary accruals computed using the Modified Jones model;

OC = the operating cycle (in days), calculated as $\frac{(AR_t + A_{t-1})/2}{(Sales/360)} + \frac{(INV_t + INV_{t-1})/2}{(COGS/360)}$;

Leverage = total liabilities (annual Compustat data item 9 plus data item 34) divided by total assets;

Book-to-Market = the market capitalization divided by the book value of common equity (annual Compustat data item 60);

Positive_DA = the value of positive discretionary accruals computed using the Modified Jones Model;

Negative_DA = the value of negative discretionary accruals computed using the Modified Jones Model;

R_CFO = the level of abnormal cash flows from operations;

R_PROD = the level of abnormal production costs, where production costs are defined as the sum of cost of goods sold and the change in inventories;

R_DISX = the level of abnormal discretionary expenses, where discretionary expenses are the sum of advertising expenses (annual Compustat data item 45), R&D expenses (annual Compustat data item 46) and SG&A expenses (annual Compustat data 189);

RM_PROXY = the sum of the standardized three real earnings management proxies, i.e., *R_CFO*, *R_PROD* and *R_DISX*;

BONUS (%) = the average bonus compensation as a proportion of total compensation received by the CEO and CFO of the firm;

BONUS ($) = the *BONUS* variable in ExecuComp in thousands of dollars;

EX_OPTIONS (%) = exercisable options, which is the number of unexercised options that the executive held at year-end that were vested scaled by total outstanding shares of the firm;

UN_OPTIONS (%) = unexercisable options defined as the number of unexercised options (excluding options grants in the current period) that the executive held at year-end that had not vested scaled by total outstanding shares of the firm;

GRNT_OPTION (%) = new option grants made during the current period scaled by total outstanding shares of the firm; and

OWNER = the sum of restricted stock grants in the current period and the aggregate number of shares held by the executive at year-end (excluding stock options) scaled by total outstanding shares of the firm.

Table 1, Panel B provides summary statistics of the ExecuComp sample.[17] Both samples are significantly larger (at the 0.001 level) in terms of total assets and market capitalization when compared to the "average" firm listed on Compustat. Sample firms have a 13 percent annual growth in sales, a market-to-book ratio of 4.94, and a leverage ratio of 0.41 (18 percent, 4.70, and 0.22, respectively, for the ExecuComp subsample).

The operating cycles of our sample firms is, on average, approximately 142.56 days (124.69 days for the ExecuComp subsample), suggesting that accruals are likely to reverse in the subsequent year. Consistent with prior research we find a positive and significant correlation (0.019, Pearson; 0.053 Spearman; both significant at the 1 percent level, results not tabulated) between the operating cycle and discretionary accruals (Dechow and Dichev 2002). Finally, requiring the availability of ExecuComp data while considerably increasing firm size does not seem to have a significant impact on fundamental measures such as leverage, growth of sales, or market-to-book ratios.

---

[17] Whenever possible, we perform the tests on both the full and the ExecuComp samples to assess the impact of the ExecuComp selection on our results.

As expected, *TA* (total accruals deflated by prior-year total assets) is negative at −0.10 (−0.07 for the ExecuComp subsample) with a standard deviation of 0.25 (0.12 for the for the ExecuComp subsample). In contrast the average *DA* (discretionary accruals) is 0.00 (standard deviation of 0.20) for the full sample and −0.01 (0.11) for the ExecuComp subsample.[18]

While the average *DA* is zero, we find that positive discretionary accruals (*Positive_DA*) are, on average, larger in magnitude than negative discretionary accruals (*Negative_DA*). This is not only true for the mean, but also for the median and the 75th percentiles. This is however not true for the ExecuComp sample (except at the 75th percentile). Thus, for the full sample, it appears that larger earnings increasing *DA*s are followed by smaller but more frequent reversals.

The main variable of interest is the absolute value of discretionary accruals (*ABS_DA*). We use the absolute value because our hypotheses do not predict any specific direction for earnings management. Moreover, the absolute value also captures accrual reversals following earnings management. The average for *ABS_DA* is 0.11 (0.07 for the ExecuComp sample). At first, this may seem a large value as percentage of total assets. However, recall that while *DA* has a mean of zero, the mean is shifted to the right by taking absolute values.[19,20]

Nevertheless, we conduct some additional analyses to get more reassurance on the magnitudes of the discretionary accrual measure. First, we selected a few firms that had very high absolute values of *DA* as percentage of total assets (higher than 15 percent) and examined their financial statements.[21] For the 12 firms that meet this criterion, we find that more than 50 percent of them had large asset impairment and restructuring charges and write-offs related to goodwill, while the rest had merger-related charges and significant growth in accounts receivables and inventory, which (at least partly) explains the large magnitudes in *DA*s that we observe. To check whether our results are affected by outliers, we winsorize the top and bottom 1 percent of the distribution. On winsorizing, the mean (median) of *ABS_DA*, is 0.06 (0.05) for the full sample and 0.04 (0.04) for the ExecuComp sample. These magnitudes of discretionary accruals appear more realistic. Moreover, our cross-sectional results and inferences are not affected by repeating all our analyses using the winsorized sample (results not tabulated).

Finally, we also check the discretionary accruals of the firms that were involved in accounting scandals. Untabulated results indicate that all but one of the firms involved

---

[18] Note that the mean Jones model residual is zero by construction (residual of a regression), whereas the mean modified Jones model residual is not constrained to be zero by construction.

[19] If X follows a normal distribution with mean μ and variance σ^2 then the expected absolute value of X is given by:

$$E[|X]] = \sigma\sqrt{2/\pi e \frac{\mu^2}{2\sigma^2}} + \mu\left[2\Phi\left(\frac{\mu}{\sigma}\right) - 1\right]$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution function.

[20] We also perform two "sanity checks." First, we generated 1,000 random variables from a normal distribution with mean 0.00 and a standard deviation of 0.20. The mean of the absolute values of those random variables is 0.173. Second, we computed the ROA and the absolute value of the ROA for our sample firms. While the average ROA for our sample firms is −0.06 with a standard deviation of 0.191 (0.046 and 0.207 for the ExecuComp subsample) the absolute value of ROA is 0.22 (0.118 for the ExecuComp subsample).

[21] We thank one of the referees for this suggestion.

in accounting scandals fall in the 90th percentile of the *ABS_DA* distribution in the SCA period.

The last three rows of Panels A and B of Table 1 report our proxies for real earnings management. Comparing the 25th and 75th percentiles of the real earnings management proxies to *DA* suggests that accrual-based earnings management takes larger values. This observation is consistent with Graham et al.'s (2005) survey, which suggests that real earnings management is more costly than accrual based earnings management.

The last five rows of Panel B of Table 1 report both the dollar amount obtained from bonuses and characteristics of equity-based compensation, i.e., exercisable and unexercisable options, new option grants, and stock ownership. For the entire sample period, the average ExecuComp CEO received annual bonuses of $454,040, representing 16 percent of their total compensation. New option grants and other unexercisable options are on average 0.13 percent and 0.21 percent of outstanding shares respectively, while exercisable options are on average 0.75 percent of outstanding shares. The sum of restricted stock grants and aggregate shares held by the CEO averages 5.19 percent.

Table 2, Panel A summarizes time-trends of the accrual and real earnings management proxies. To summarize the data, we regress each of the variables on a time-trend variable, *Time*, defined as the difference between the year and 1987, and two dummy variables, *SCA*, which takes the value of 1 in the scandal period (years 2000 and 2001), and 0 otherwise, and *SOX*, which takes the value of 1 in the post-SOX period (years 2002, 2003, and 2005), and 0 otherwise. We choose this procedure to describe the variables because many of our variables exhibit significant time trends (nonstationarity), rendering traditional summary statistics uninformative.

The coefficient for the time trend in Table 2, Panel A, row 1 ($\hat{b}$) indicates that the magnitude of discretionary accruals (*ABS_DA*) has been increasing significantly (at the 1 percent level) over the sample period with that increase being nearly symmetric for both *Positive_DA* and *Negative_DA* (rows 2 and 3).

The magnitude of discretionary accruals increased significantly in the SCA period ($\hat{c}$ in Row 1) with positive (i.e., income-increasing, Row 2) discretionary accruals contributing twice as much to that increase than income-decreasing *DA*s (Row 3, both significant at the 1 percent level).

Finally, the magnitude of discretionary accruals declined significantly in the post-SOX period ($\hat{d}$ in Row 1). Moreover, the magnitude (absolute value) of the coefficient for *Positive_DA* ($\hat{d}$ in Row 2) is approximately three times larger than the coefficient for *Negative_DA* ($\hat{d}$ in Row 3) suggesting that most of that decline in accrual-based management results from the reduction of *Positive_DA*s.

Figures 2 and 3, Panels A and B provide graphical illustrations of these results. Figure 2 indicates that the SCA period was, indeed, associated with a high level of earnings management. Figure 3, Panels A and B plot the trends in positive and negative discretionary accruals. Positive discretionary accruals peaked in the SCA period and negative discretionary accruals were the lowest in that period. These trends reversed in the post-SOX period.

Among the real earnings management variable, except for *R_DISX*, which increased over the period, the other real earnings management variables do not show an increasing trend over the sample period (the time trend coefficients are either not significant or negative). Abnormal production costs and abnormal cash flows both increased significantly in the post-SOX period, but were not significantly higher in the SCA period. On the other hand, abnormal discretionary expenses were significantly higher in the SCA period and significantly lower in the post-SOX period. The combined variable *RM_PROXY* shows a

TABLE 2
Earnings Management Metrics: Time Trends and Correlation Matrix

Panel A: Time Trend Analysis of Earnings Management Metrics over Time, 1987–2005

Dependent	$Dep_{jq} = a + b \times Time + c \times SCA + d \times SOX$				
Variables	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{d}$	**Adjusted R²**
ABS_DA	0.053***	0.003***	0.027***	−0.015***	0.024
Positive_DA	0.028***	0.002***	0.026***	−0.019**	0.029
Negative_DA	−0.040***	−0.003***	−0.012***	0.006***	0.015
R_PROD	−0.040***	−0.002	0.008	0.032***	0.009
R_CFO	0.022***	−0.003**	−0.007***	0.007**	0.007
R_DISX	0.028***	0.005***	0.008***	−0.011***	0.010
RM_PROXY	−0.018***	−0.005**	0.033**	0.041***	0.011

Panel B: Correlation between the Earnings Management Proxies, 1987–2005

	DA	*ABS_DA*	*R_CFO*	*R_PROD*	*R_DISX*	*RM_PROXY*
DA	1	0.398***	−0.237***	−0.029***	−0.163***	−0.331***
ABS_DA	0.298***	1	−0.184***	−0.047***	−0.174***	−0.153***
R_CFO	−0.214***	−0.114***	1	−0.274***	−0.142***	0.447***
R_PROD	−0.024***	−0.314***	−0.183***	1	−0.243***	0.579***
R_DISX	−0.143***	−0.164***	−0.214***	−0.187***	1	0.487***
RM_PROXY	−0.237***	−0.149***	0.381***	0.427**	0.422***	1

---

*, **, *** Significant at 10 percent, 5 percent, and 1 percent levels, respectively.

Variable Definitions:

DA = discretionary accruals computed using the Modified Jones Model;
ABS_DA = the absolute value of discretionary accruals computed using the Modified Jones Model;
Positive_DA = the value of positive discretionary accruals computed using the Modified Jones Model;
Negative_DA = the value of negative discretionary accruals computed using the Modified Jones Model;
TIME = a trend variable equal to the difference between the current year and 1987;
SCA = a dummy variable equal to 1 if the year is either 2000 or 2001; and
SOX = a dummy variable for years equal to 2002, 2003, 2004, and 2005.

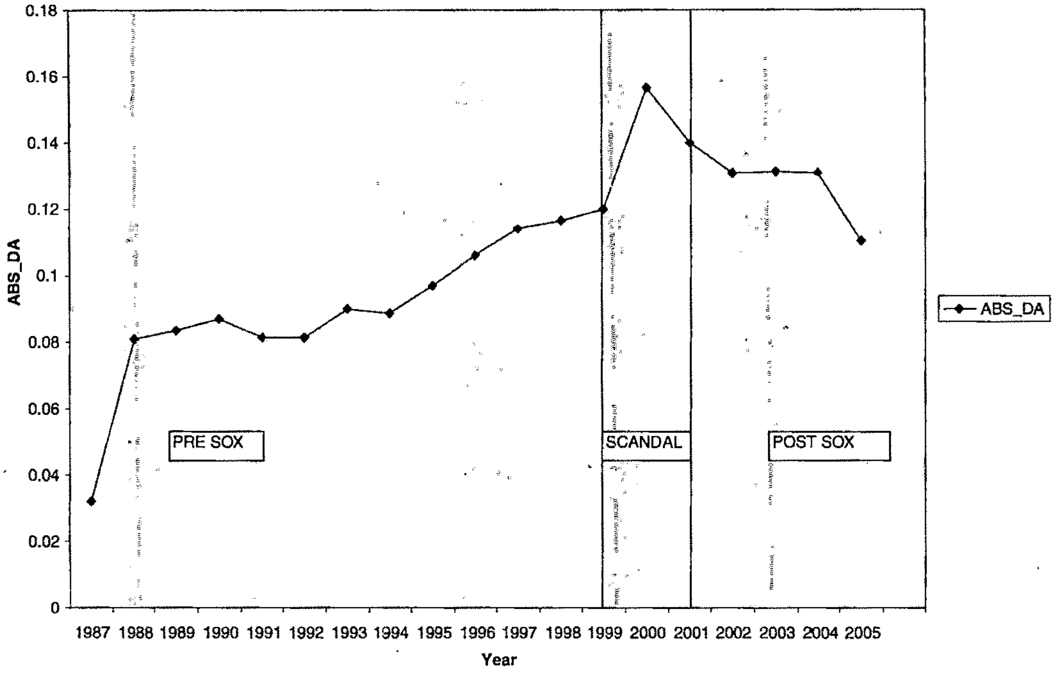
decreasing trend over time, but is significantly higher both in the SCA period and in the post-SOX periods.

Figure 4 graphically illustrates these trends. A comparison of the post-SOX coefficients for accrual-based and the real earnings management suggests that there may have been a substitution effect: while accrual-based earnings management decreased, overall firms increased the use of real earnings management methods.

The correlation among accrual-based and real earnings management is reported in Table 2, Panel B. We find a significant negative relation between discretionary accruals and the real earnings management metrics, suggesting that firms appear to use these two earnings management methods as substitutes. Further, the three real earnings management variables are also negatively correlated, indicating that firms switch between real earnings management methods.

In summary, the above analysis indicates that the overall level of accrual-based earnings management decreased from the SCA period to the post-SOX period, while overall the level

FIGURE 2
Absolute Value of Discretionary Accruals over Time, 1987–2005



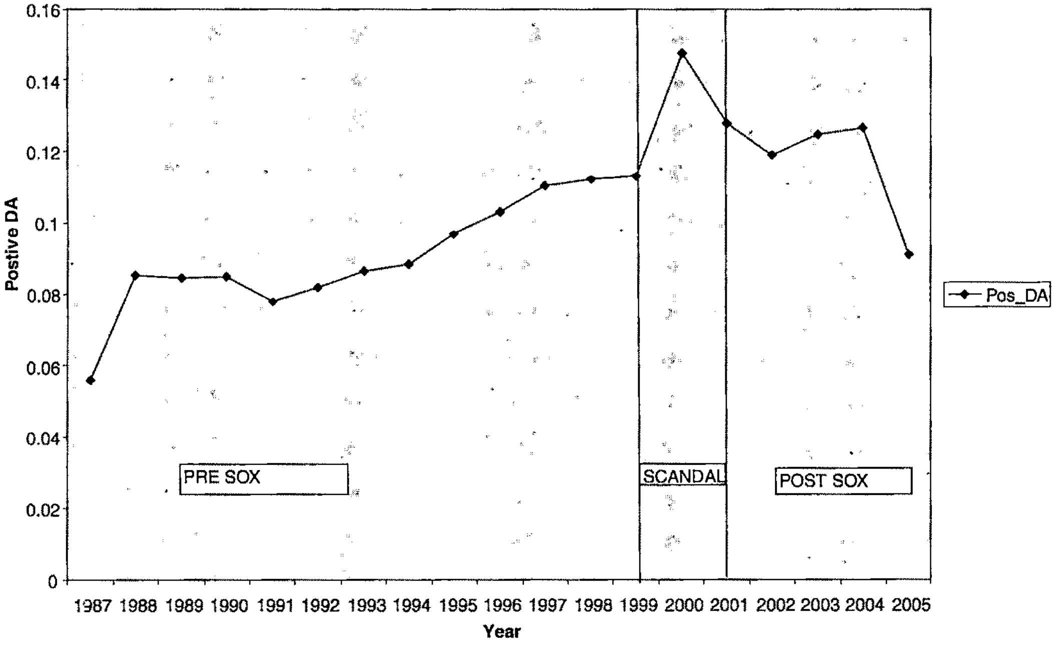
This figure plots the absolute value of discretionary accruals computed using the Modified Jones Model over the 1987–2005 sample period.

of real earnings management increased in the post-SOX period. However, there was significantly higher earnings management, particularly income-increasing earnings management, during the SCA period as compared to the pre-SCA period. One interpretation of this result is that the SCA period was characterized by higher earnings management; and the scandal firms were not just a "few bad apples," but evidence of the generally high level of accrual management. Another observation is that, although earnings management using accrual-based means increased from the pre-SCA to the SCA period, it declined significantly from the SCA period to the post-SOX period. It should be noted that most of the decline in accrual-based earnings management seems to have been due to a reduction in positive discretionary accruals.

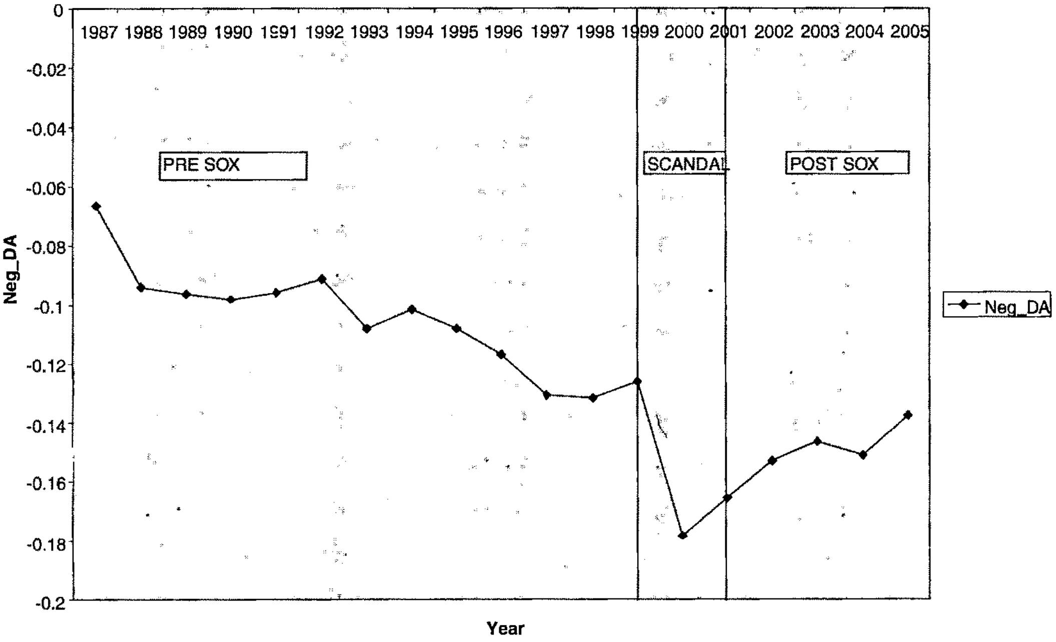
Whether this decline is caused by the passage of SOX or other concurrent events (such as the negative publicity of the most egregious governance failures and the highly visible enforcement actions directed at the offending corporate officers) cannot be inferred from this analysis, however. Moreover, firms appeared to have switched to managing earnings using real management techniques after SOX, probably because such methods of manipulation are harder to detect. How this substitution between the different types of earnings management affected the overall level of earnings management post-SOX is unclear. We formally examine the determinants of earnings management in our multivariate analysis next.

FIGURE 3
Discretionary Accruals over Time, 1987–2005

Panel A: Positive Discretionary Accruals over Time, 1987–2005



Panel B: Negative Discretionary Accruals over Time, 1987–2005



This figure plots positive and negative discretionary accruals computed using the Modified Jones Model over the 1987–2005 sample period.

FIGURE 4
Real Earnings Management Proxies over Time, 1987–2005

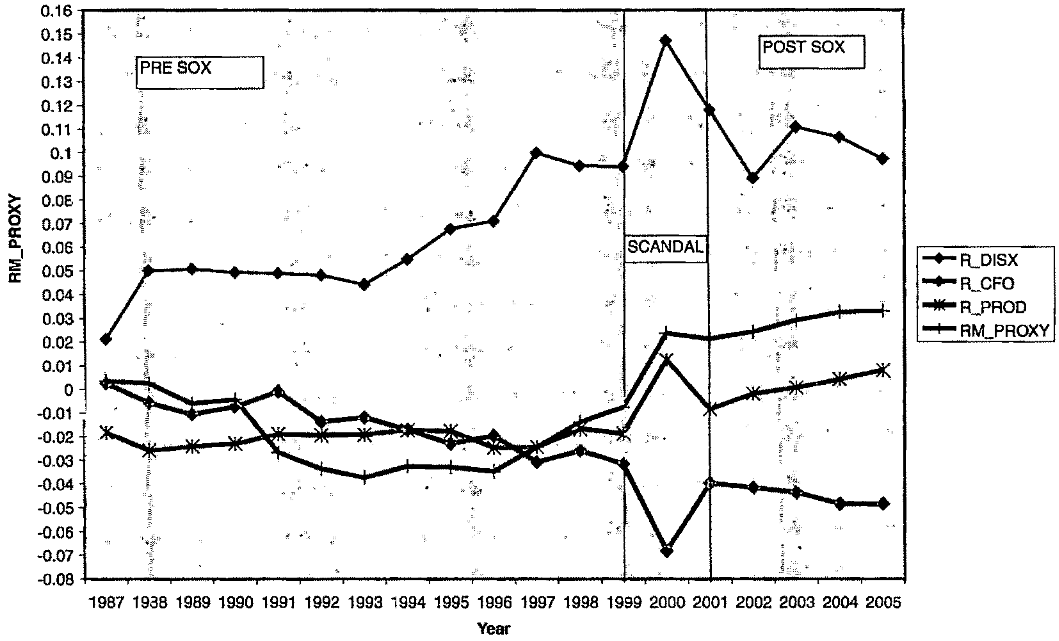


Figure 4 plots abnormal cash from operations, abnormal production costs, abnormal discretionary expenses, and the sum of the standardized three real earnings management proxies, *RM_PROXY* over the 1987–2005 sample period.

Trends In and Determinants of Earnings Management

We examine the trends in and determinants of the level of earnings management over time by estimating the following regression:

$$\begin{aligned} DEP_j = {} & \alpha_0 + \alpha_1 \times BIG_j + \alpha_2 \times \Delta GDP_j + \alpha_3 \times MKTVAL_j + \alpha_4 \times Time \\ & + \alpha_5 \times SCA + \alpha_6 \times SOX + \alpha_7 \times RM\_PROXY + \alpha_8 \times BONUS_j \\ & + \alpha_9 \times BONUS_j \times SCA + \alpha_{10} \times BONUS_j \times SOX \\ & + \alpha_{11} \times UN\_OPTION_j + \alpha_{12} \times UN\_OPTION_j \times SCA \\ & + \alpha_{13} \times UN\_OPTION_j \times SOX + \alpha_{14} \times GRNT\_OPTION_j \\ & + \alpha_{15} \times GRNT\_OPTION_j \times SCA + \alpha_{16} \times GRNT\_OPTION_j \times SOX \\ & + \alpha_{17} \times EX\_OPTION_j + \alpha_{18} \times EX\_OPTION_j \times SCA \\ & + \alpha_{19} \times EX\_OPTION_j \times SOX + \alpha_{20} \times OWNER_j \\ & + \alpha_{21} \times OWNER_j \times SCA + \alpha_{22} \times OWNER_j \times SOX \end{aligned} \tag{10}$$

where:

DEP_j = the various earnings management metrics, including discretionary accruals, positive discretionary accruals, and negative discretionary accruals;
BIG = a dummy variable equal to 1 if the auditor is a Big 5 audit firm (or their successors);
ΔGDP = the change in the GDP;
MKTVAL = market value of equity;
Time = the calendar year minus 1987;
SCA = a dummy variable that is equal to 1 for the years 2000 and 2001, and 0 otherwise (represents the SCA period);
SOX = a dummy variable that is equal to 1 for the years 2002 onward, and 0 otherwise (represents the post-SOX period);
BONUS = the average bonus compensation as a proportion of total compensation received by the CEO and the CFO of a firm;
EX_OPTION = exercisable options and is the number of unexercised options that the executives held at year-end that were vested scaled by total outstanding shares of the firm;
UN_OPTION = unexercisable options defined as the number of unexercised options (excluding option grants in the current period) that the executives held at year-end that have not vested scaled by total outstanding shares of the firm;
GRNT_OPTION = new option grants made during the current period scaled by total outstanding shares of the firm; and
OWNER = the sum of restricted stock grants in the current period and the aggregate number of shares held by the executives at year-end (excluding stock options) scaled by total outstanding shares of the firm.[22]

The compensation variables proxy for performance-based compensation and are defined in accordance with prior studies (Cheng and Warfield 2005). In addition, we separate out new options granted in the current period from other unexercisable options because current grants are likely to provide incentives to reduce earnings to benefit from lower exercise prices, which are opposite to the incentives provided by other unexercisable options. We include the variable *GRNT_OPTION* in order to capture the differential incentive effects of new option grants awarded during the current period. We include these variables to test the conjecture that the bonus and equity components of executive compensation are likely to induce opportunistic behavior by managers.[23]

We include the variable ΔGDP as a proxy for real economic activity. We include this to control for the effect of economic activity on earnings management, since what might be classified as opportunistic earnings management may, in fact, be a consequence of changing economic conditions. Discretionary accruals may also reflect firms' responses to and representations of changes in economic conditions. If this were true, then changes in earnings management metrics would coincide with changes in measures of economic activity such as operating cash flows, revenues, prior stock returns, industry performance, changes in gross domestic product, etc. Further, after controlling for changes in economic activities, there should be no relation between increases in earnings management and the compensation variables.[24]

---

[22] We also repeat the analyses by measuring the compensation related variables for the top five executives of a firm, and the results are qualitatively unchanged.

[23] The timing of variable measurement is consistent with Cheng and Warfield (2005, 448). The equity incentive variables are measured during or at the end of fiscal year *t*, whereas the earnings management variable is measured based on information disclosed after the end of fiscal year *t*.

We include control variables for the auditors in the above regression to examine whether the earnings management activity of firms audited by the large audit firms were different from the rest of the sample firms over the three subperiods analyzed. Note that we make no claim that differences in the earnings management activities (if any) of these firms were the result of the monitoring activities of the audit firms, since there could be a self-selection by certain types of firms in their selection of big audit firms. In addition, to the extent that audit firms specialize in specific industries and levels of earnings management are likely to vary across industries, the audit firm dummies may also control for industry characteristics. Finally, we include the market value of equity as a proxy for firm size.

Next, we examine the trends in and determinants of real earnings management activities of firms by estimating the following regression:

$$\begin{aligned} DEP_j = {} & \alpha_0 + \alpha_1 \times BIG_j + \alpha_2 \times \Delta GDP_j + \alpha_3 \times MKTVAL_j + \alpha_4 \times Time \\ & + \alpha_5 \times SCA + \alpha_6 \times SOX + \alpha_7 \times ABS\_DA_j + \alpha_8 \times BONUS_j \\ & + \alpha_9 \times BONUS_j \times SCA + \alpha_{10} \times BONUS_j \times SOX \\ & + \alpha_{11} \times UN\_OPTION_j + \alpha_{12} \times UN\_OPTION_j \times SCA \\ & + \alpha_{13} \times UN\_OPTION_j \times SOX + \alpha_{14} \times GRNT\_OPTION_j \\ & + \alpha_{15} \times GRNT\_OPTION_j \times SCA + \alpha_{16} \times GRNT\_OPTION_j \times SOX \\ & + \alpha_{17} \times EX\_OPTION_j + \alpha_{18} \times EX\_OPTION_j \times SCA \\ & + \alpha_{19} \times EX\_OPTION_j \times SOX + \alpha_{20} \times OWNER_j \\ & + \alpha_{21} \times OWNER_j \times SCA + \alpha_{22} \times OWNER_j \times SOX \end{aligned} \tag{11}$$

where DEP_j represents the three real earnings management metrics, R_CFO, R_PROD, and R_DISX, and the combined variable RM_PROXY. All other variables are as defined above.

Firms may follow an overall earnings management strategy and use a mix of real and accrual-based earnings management tools. Alternatively, they can choose between the two management techniques, using the technique that is less costly for them. To control for this possibility, we also include a variable representing accrual-based earnings management (we include *ABS_DA* in the regression). If it were more costly for firms to manage earnings using accrual-based techniques after the passage of SOX and they substituted this by using real management techniques instead, then we should observe a significant increase in the latter after SOX. On the other hand, firms could have decreased earnings management activities as a whole—in which case we would observe a decrease in real earnings management after SOX as well. The results of these regressions are discussed next.

---

[24] In the "SUSPECT Firms Analysis" section we discuss some robustness tests we conduct to further test whether we are indeed capturing earnings management activities.

Results

Tables 3 and 4 present the results of the determinants of the level of earnings management by firms. Table 3 reports the results when we use the absolute value of discretionary accruals, *ABS_DA*, and when we split discretionary accruals into positive and negative discretionary accruals. Table 4 reports the results when we use the measures for real earnings management activities. We discuss the results for the accrual-based earnings management variables first.

Consistent with the preliminary analysis reported in Table 2, we find a positive trend in the level of earnings management, including income-increasing earnings management, and a negative trend in income-decreasing earnings management. This indicates that overall earnings management increased over the sample period. The dummy variable *SOX* is negative and significant for *ABS_DA* as well as for positive discretionary accruals and positive and significant for negative discretionary accruals. This suggests that, controlling for the other independent variables, the period after SOX was characterized by lower accrual-based earnings management and this decrease primarily resulted from a decrease in income-increasing earnings management. Several simultaneous occurrences could have contributed to a decrease in earnings management activities after passage of SOX, including the increased vigilance of investors, auditors and regulators, and greater care taken by managers in financial reporting after the adverse publicity caused by the scandals. Thus, we are cautious in attributing the decrease in the level of earnings management solely to the passage of SOX from this analysis.

Unlike our analysis in Table 2, the dummy variable representing the scandal period is not significant for any of the variables. This implies that the entire increase in the accrual-based earnings management in the SCA period is related to the increase in equity-based compensation. Further, we find that the percentage of bonus compensation is not correlated with earnings management for the entire period, and this association did not significantly change in the SCA or the post-SOX periods. Positive discretionary accruals, however, were significant and positively associated with the percentage of compensation received from bonuses in the scandal period.

Consistent with our conjecture, the percentage of compensation derived from unexercised options and stock ownership is significantly positively associated with discretionary accruals as well as with positive discretionary accruals. For unexercised options this effect increased significantly in the SCA period for positive discretionary accruals and decreased significantly in the post-SOX period. This suggests that option compensation increased managers' incentives to manipulate earnings upward (e.g., Fuller and Jensen 2002; Greenspan 2002; Coffee 2003), and this effect was significantly higher during the period surrounding the corporate scandals. One possible explanation for the decline after SOX could be the penalties on incentive compensation introduced by SOX. Unexercised options and ownership are significantly negatively correlated with negative discretionary accruals overall and after SOX. This suggests that the presence of stock and options in the compensation structure reduces the incentive to manage earnings downward. Unexercised options are also significantly negatively related to negative discretionary accruals in the scandal period indicating an even lesser inclination toward managing earnings downward during this period.

New option grants (*GRNT_OPTION*) are positively associated with *ABS_DA* overall, not significantly related to *ABS_DA* during the scandal period and significantly negatively related to *ABS_DA* in the post-SOX period. The results for the positive and negative discretionary accruals are more interesting and are consistent with the notion that managers

TABLE 3
Determinants of Accrual Based Earnings Management Activities 1987–2005

$$\begin{aligned} DEP_j = {} & \alpha_0 + \alpha_1 \times BIG_j + \alpha_2 \times \Delta GDP_j + \alpha_3 \times MKTVAL_j + \alpha_4 \times Time \\ & + \alpha_5 \times SCA + \alpha_6 \times SOX + \alpha_7 \times RM\_PROXY + \alpha_8 \times BONUS_j \\ & + \alpha_9 \times BONUS_j \times SCA + \alpha_{10} \times BONUS_j \times SOX + \alpha_{11} \times UN\_OPTION_j \\ & + \alpha_{12} \times UN\_OPTION_j \times SCA + \alpha_{13} \times UN\_OPTION_j \times SOX \\ & + \alpha_{14} \times GRNT\_OPTION_j + \alpha_{15} \times GRNT\_OPTION_j \times SCA \\ & + \alpha_{16} \times GRNT\_OPTION_j \times SOX + \alpha_{17} \times EX\_OPTION_j \\ & + \alpha_{18} \times EX\_OPTION_j \times SCA + \alpha_{19} \times EX\_OPTION_j \times SOX \\ & + \alpha_{20} \times OWNER_j + \alpha_{21} \times OWNER_j \times SCA + \alpha_{22} \times OWNER_j \times SOX \end{aligned}$$

	ABS_DA		*Positive_DA*		*Negative_DA*	
	Coefficient	**(t-stat)**	**Coefficient**	**(t-stat)**	**Coefficient**	**(t-stat)**
Intercept	0.118	(18.87)	0.022	(9.67)	−0.049	(−14.17)
BIG	−0.027	(−5.87)	−0.007	(−2.32)	−0.017	(−2.79)
ΔGDP	−0.114	(−6.21)	−0.141	(−5.68)	0.098	(2.99)
MKTVAL	−0.007	(−17.36)	−0.002	(−7.76)	0.007	(11.67)
Time	0.003	(8.21)	0.003	(13.87)	−0.012	(−2.71)
SCA	−0.002	(−0.42)	−0.001	(−0.41)	−0.002	(−0.49)
SOX	−0.019	(−4.32)	−0.024	(−6.76)	0.008	(3.44)
RM_PROXY	−0.427	(−5.81)	−0.529	(−6.17)	0.112	(1.63)
BONUS	0.002	(0.48)	−0.003	(−1.24)	0.003	(0.81)
BONUS × *SCA*	−0.005	(−0.43)	0.027	(2.53)	0.016	(0.79)
BONUS × *SOX*	0.014	(1.27)	0.006	(1.07)	0.003	(0.51)
UN_OPTION	0.423	(10.97)	0.467	(6.97)	−0.219	(−6.97)
UN_OPTION × *SCA*	−0.124	(−1.21)	0.137	(5.09)	−0.086	(−1.89)
UN_OPTION × *SOX*	−0.237	(−3.94)	−0.124	(−6.14)	−0.046	(−1.59)
GRNT_OPTION	0.119	(2.93)	−0.009	(−2.86)	0.062	(2.23)
GRNT_OPTION × *SCA*	−0.087	(−0.91)	−0.117	(−2.04)	0.019	(2.39)
GRNT_OPTION × *SOX*	−0.074	(−3.63)	−0.009	(−1.09)	0.008	(1.53)
EX_OPTION	0.121	(1.12)	0.081	(0.87)	−0.052	(−0.46)
EX_OPTION × *SCA*	−0.006	(−0.74)	−0.042	(−1.08)	0.003	(0.73)
EX_OPTION × *SOX*	−0.009	(−1.09)	−0.007	(−1.18)	−0.002	(−0.71)
OWNER	0.092	(5.58)	0.096	(4.83)	−0.048	(−3.37)
OWNER × *SCA*	0.004	(0.49)	0.007	(1.19)	0.002	(0.44)
OWNER × *SOX*	−0.033	(−3.84)	−0.057	(−5.43)	−0.016	(−2.27)
Adjusted R^2	0.084		0.081		0.061	
F-value (Pr > F)	169.16 (<0.0001)		176.34 (<0.0001)		163.21 (<0.0001)	

Variable Definitions:
The dependent variable is either *ABS_DA*, *Positive_DA* or *Negative_DA* where *ABS_DA* is the absolute value of discretionary accruals computed using the Modified Jones Model, *Positive_DA* is the value of positive discretionary accruals computed using the Modified Jones model and *Negative_DA* is the value of negative discretionary accruals computed using the Modified Jones model.

(*continued on next page*)

TABLE 3 (continued)

RM_PROXY = the sum of the standardized three real earnings management proxies, i.e., *R_CFO, R_PROD* and *R_DISX*.
ΔGDP = the percent change in the real gross domestic product from the previous year;
BIG = a dummy variable if the company is audited by a Big 5 (4) auditor;
MKTVAL = market value of equity;
TIME = a trend variable equal to the difference between the current year of observation and 1987;
SCA = a dummy variable equal to 1 if the year of observation is in 2000 or 2001;
SOX = a dummy variable equal to 1 if the year of observation is in 2002, 2003, 2004, or 2005;
BONUS = the average bonus compensation as a proportion of total compensation received by the CEO and CFO of the firm;
EX_OPTION = the number of unexercised options that the executive held at year-end that were vested scaled by total outstanding shares of the firm;
UN_OPTION = the number of unexercised options (excluding options grants in the current period) that the executive held at year-end that had not vested scaled by total outstanding shares of the firm;
GRNT_OPTION = new option grants made during the current period scaled by total outstanding shares of the firm; and
OWNER = the sum of restricted stock grants in the current period and the aggregate number of shares held by the executive at year-end (excluding stock options) scaled by total outstanding shares of the firm.

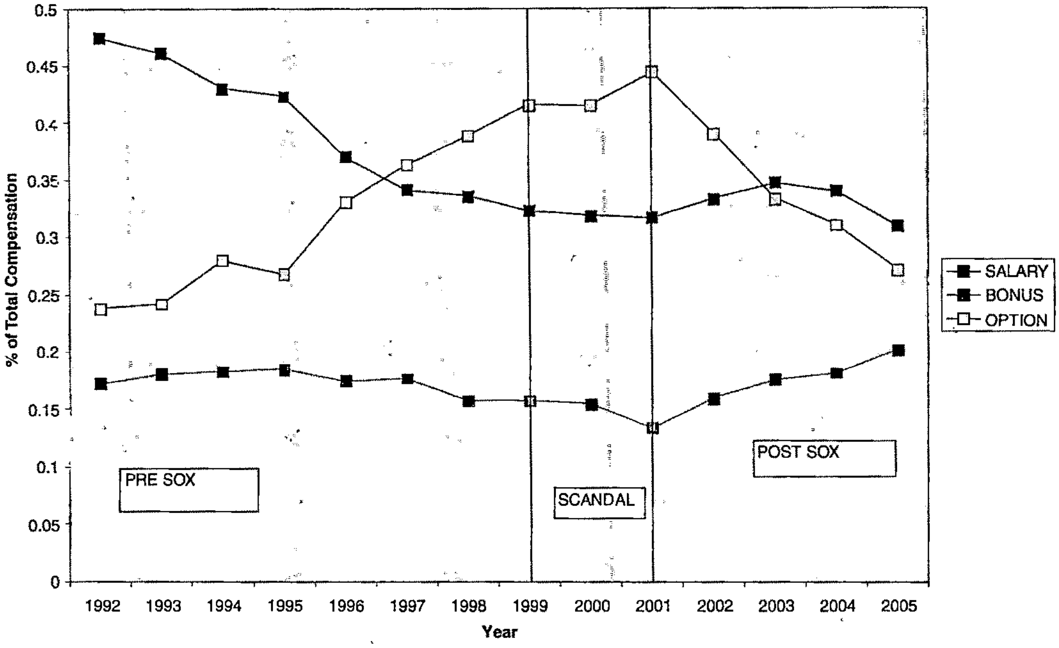
The t-statistics are calculated using Huber-White standard errors.

manage earnings downward during periods when options are granted to benefit from lower exercise prices. Specifically, new option grants are significantly negatively related to positive discretionary accruals, and continue to be significant and negative in the SCA period. On the other hand, new option grants are significantly positively related to negative discretionary accruals, and continue to be significant and positive in the SCA period. However, the relation between new option grants and positive or negative discretionary accruals is not significant after SOX.

The coefficient on exercisable options is positive (for *ABS_DA* and *Positive_DA*) and negative (for *Negative_DA*) but statistically insignificant, a result consistent with the evidence in Cheng and Warfield (2005). This finding is somewhat unintuitive because one would expect both unexercisable and exercisable options to have similar effects on managers' incentives to manage earnings. Cheng and Warfield (2005) argue that equity incentives lead to earnings management via future trading in the company's stock, and because exercisable options are not correlated with future trading (unlike unexercisable options and ownership), these are not related to earnings management. This is likely to explain our results as well, but we do not examine this reasoning further as this is beyond the scope of our analysis.

These results are reinforced by Figure 5, which plots bonus and option compensation (based on a Black-Scholes valuation) as a percentage of total compensation. The figure indicates that while bonus compensation was relatively stable around 15–20 percent over the entire period, stock option compensation increased over time and peaked to as much as 45 percent of total compensation in 2001 and then gradually decreased back to the 25 percent range by 2005. This increase in option-based compensation during the scandal period provides an explanation for the significantly higher association between options and earnings management during that period. This is consistent with the evidence in Cheng and Warfield (2005) and Bergstresser and Philippon (2006). Overall, the evidence in Cheng and Warfield (2005) suggests that equity incentives derived from stock option compensation lead to incentives to engage in accrual-based earnings management activities. In addition, Bergstresser and Philippon (2006) show that the use of discretionary accruals to manipulate reported earnings is more pronounced at firms where the CEO's compensation is more closely tied to stock options. The evidence we document also supports the assertions made

FIGURE 5
Compensation over Time, 1992–2005



This figures plots trends in executive compensation (for CEOs and CFOs) over the sample period, 1992–2005. All variables are presented as a percentage of total compensation.

by Coffee (2003), Greenspan (2002), and Fuller and Jensen (2002), among others, that stock-based compensation increase managers' incentives to inflate reported earnings and, consequently, stock prices.

Overall, these results identify compensation, primarily derived from unexercised stock options, as being significantly correlated with accrual-based earnings management in the period leading up to the passage of SOX. The association between options and accrual-based earnings management increased during the scandal period and decreased in the post-SOX period. As mentioned earlier, one explanation for the latter result could be the provisions in Sections 302 and 304 in SOX that mandate the return of any incentive compensation owing to material noncompliance with any financial reporting requirement. This could have made executives with stock options more conservative in their financial reporting practices.

The coefficient for the change in GDP (ΔGDP), our proxy for real economic activity, is negative and significant for discretionary accruals and for positive discretionary accruals and positive and significant for negative discretionary accruals.[25] This result suggests that after controlling for managers' opportunistic incentives, poor economic conditions are accompanied by more discretionary accruals. However, even after controlling for economic conditions, we find a significant reduction in earnings management activities following the passage of SOX. We find evidence that earnings management is significantly less for larger firms and for clients of big audit firms.[26]

---

[25] The documented results are robust to an alternative measure of economic activities, namely, industry-adjusted ROA. We also repeated the analysis by including measures of growth, either the book-to-market ratio or the growth in sales as an additional control variable, and our results are qualitatively unaltered by these inclusions. Including prior-year stock returns as well as using the logarithm of market value of equity as alternative measures of size does not affect the reported results.

The results for the real earnings management variables (Table 4) are consistent across all three real earnings management variables as well as for the combined variable *RM_PROXY*; therefore, we report all three sets of results but discuss those for only the combined variable *RM_PROXY* (we point out differences in the results whenever applicable). Consistent with the results in Table 2, the level of overall real earnings management decreased over the sample period, but increased significantly during the scandal period and after SOX. However, discretionary spending increased over the sample period and in the scandal period, and decreased significantly after SOX.

The coefficient estimate on accrual-based earnings management variable, *ABS_DA*, is negative and significant, which indicates that firms substitute between real and accrual-based earnings management activities, a result consistent with Graham et al. (2005) and Zang (2006).

The coefficient estimates on bonus, unexercisable options, and stock ownership are negative and significant, indicating that executives with higher bonus, unexercisable options, and stock ownership managed earnings using real techniques less. However, except for discretionary spending, the interaction terms between unexercisable options and SOX and between ownership and SOX are positive and significant, suggesting that executives with more unexercisable options and stock managed earnings using real methods more in the post-SOX period. New option grants and exercisable options are not significant for any of the real earnings management variables.

In light of the results obtained for the accrual-based earnings management variables that indicate that executives with more stock and options manage earnings more, the above findings could result from the substitution between real and accrual-based earnings management activities undertaken by firms. In the post-SOX period, we found earlier that those individuals with more stock and unexercisable options were less likely to use accrual-based means to manage earnings. These results support the conjecture that after SOX, stock and options provided increased incentives to manage earnings using real methods. In other words, if managers expect earnings management using real methods to be harder to detect, the passage of SOX is likely to result in a substitution of real earnings management for accrual-based earnings management. The results for the other control variables are similar to those obtained in the case of the accrual-based earnings management variables.

Finally, in order to allow for production differences between manufacturers and retailers, we refined our real earnings management analysis by distinguishing between manufacturing firms and non-manufacturing firms. Based on two-digit SIC codes, we divided the sample firms into manufacturing and non-manufacturing industries and repeated the analysis. We continue to find for non-manufacturing firms results that are consistent with those documented in column 2 of Table 4. We tested an alternative specification in which we included a dummy variable for manufacturing firms in Table 4 and interacted it with the time trend variables. We defined this dummy to take the value of 1 if the industry's two-digit SIC code falls between 20 and 39. The dummy variable came in positive and significant (an intercept effect), and the interaction with *SCA*, although negative, was insignificant, whereas the interaction with *SOX* was significantly positive, consistent with the results

---

[26] In unreported results, we did not find significant differences in earnings management between Big and non-Big audit firm clients between the pre-SCA, the SCA, and the post-SOX periods.

TABLE 4
Determinants of Real Earnings Management Activities 1987–2005

$$\begin{aligned} DEP_j = {} & \alpha_0 + \alpha_1 \times BIG_j + \alpha_2 \times \Delta GDP_j + \alpha_3 \times MKTVAL_j + \alpha_4 \times Time \\ & + \alpha_5 \times SCA + \alpha_6 \times SOX + \alpha_7 \times ABS\_DA_j + \alpha_8 \times BONUS_j \\ & + \alpha_9 \times BONUS_j \times SCA + \alpha_{10} \times BONUS_j \times SOX + \alpha_{11} \times UN\_OPTION_j \\ & + \alpha_{12} \times UN\_OPTION_j \times SCA + \alpha_{13} \times UN\_OPTION_j \times SOX \\ & + \alpha_{14} \times GRNT\_OPTION_j + \alpha_{15} \times GRNT\_OPTION_j \times SCA \\ & + \alpha_{16} \times GRNT\_OPTION_j \times SOX + \alpha_{17} \times EX\_OPTION_j \\ & + \alpha_{18} \times EX\_OPTION_j \times SCA + \alpha_{19} \times EX\_OPTION_j \times SOX \\ & + \alpha_{20} \times OWNER_j + \alpha_{21} \times OWNER_j \times SCA + \alpha_{22} \times OWNER_j \times SOX \end{aligned}$$

	R_CFO		***R_PROD***		***R_DISX***		***RM_PROXY***	
	Coefficient (t-stat)		**Coefficient (t-stat)**		**Coefficient (t-stat)**		**Coefficient (t-stat)**	
Intercept	0.014	(3.24)	0.075	(9.34)	0.012	(5.41)	−0.039	(−3.67)
BIG	−0.005	(−0.75)	−0.004	(−0.89)	−0.008	(−0.49)	−0.003	(−0.74)
ΔGDP	−0.157	(−2.58)	−0.194	(−4.09)	−0.186	(−4.73)	−0.167	(−4.93)
MKTVAL	0.009	(2.07)	−0.005	(−4.29)	0.008	(5.96)	−0.032	(−5.41)
Time	−0.002	(−2.21)	−0.002	(−1.61)	0.004	(2.74)	−0.004	(−2.89)
SCA	−0.006	(−1.13)	−0.014	(−1.76)	0.008	(3.76)	0.033	(3.17)
SOX	0.009	(1.93)	0.042	(5.07)	−0.018	(−3.77)	0.034	(5.71)
ABS_DA	−0.087	(−8.53)	−0.097	(−8.64)	−0.017	(−1.81)	−0.083	(−6.16)
BONUS	−0.042	(−5.12)	−0.057	(−7.82)	−0.023	(−1.91)	−0.042	(−4.14)
BONUS × *SCA*	0.005	(0.91)	0.004	(0.63)	0.009	(1.12)	0.003	(0.96)
BONUS × *SOX*	0.008	(1.09)	0.003	(0.78)	0.003	(0.68)	0.004	(0.82)
UN_OPTION	−0.173	(−4.38)	−0.163	(−8.07)	−0.117	(−7.95)	−0.154	(−7.06)
UN_OPTION × *SCA*	0.054	(1.99)	0.008	(1.93)	−0.009	(−1.09)	0.042	(1.79)
UN_OPTION × *SOX*	0.107	(5.34)	0.064	(4.17)	−0.013	(−3.17)	0.084	(3.52)
GRNT_OPTION	0.006	(0.96)	0.005	(0.37)	−0.002	(−0.31)	0.004	(1.03)
GRNT_OPTION × *SCA*	0.004	(1.34)	0.001	(0.67)	−0.006	(−1.04)	0.003	(1.27)
GRNT_OPTION × *SOX*	0.003	(0.69)	0.004	(1.16)	−0.001	(0.39)	0.002	(0.59)
EX_OPTION	−0.173	(−1.51)	−0.072	(−0.58)	−0.072	(−0.83)	−0.141	(−0.68)
EX_OPTION × *SCA*	−0.041	(−0.89)	−0.022	(−0.74)	−0.004	(−0.89)	−0.041	(−0.77)
EX_OPTION × *SOX*	0.006	(0.53)	0.004	(0.31)	0.004	(0.27)	0.004	(0.52)
OWNER	−0.243	(−3.93)	−0.284	(−2.93)	−0.149	(−2.68)	−0.177	(−2.79)
OWNER × *SCA*	0.129	(2.53)	0.179	(2.81)	−0.044	(−2.29)	0.121	(2.83)
OWNER × *SOX*	0.189	(4.72)	0.211	(3.31)	−0.008	(−1.97)	0.183	(3.19)
Adjusted R^2	0.649		0.734		0.756		0.791	
F-value (Pr > F)	2716.57 (<0.0001)		2527.96 (<0.0001)		2591.26 (<0.0001)		2726.39 (<0.0001)	

(*continued on next page*)

TABLE 4 (continued)

Variable Definitions:

The dependent variable is either *R_CFO, R_PROD, R_DISX* or *RM_PROXY* where *R_CFO* represents the level of abnormal cash flows from operations, *R_PROD* represents the level of abnormal production costs, where production costs are defined as the sum of cost of goods sold and the change in inventories and *R_DISX* represents the level of abnormal discretionary expenses, where discretionary expenses are the sum of advertising expenses, R&D expenses and SG&A expenses, and *RM_PROXY* is the sum of the standardized three real earnings management proxies, i.e., *R_CFO, R_PROD* and *R_DISX.*

ΔGDP = the percent change in the real gross domestic product from the previous year;
BIG = a dummy variable if the company is audited by a Big 5 (4) auditor;
MKTVAL = market value of equity;
TIME = a trend variable equal to the difference between the year of observation and 1987;
SCA = a dummy variable equal to 1 if the year of observation is in 2000 or 2001;
SOX = a dummy variable equal to 1 if the year of observation is in 2002, 2003, 2004, or 2005;
ABS_DA = the absolute value of discretionary accruals computed using the Modified Jones model;
BONUS = the average bonus compensation as a proportion of total compensation received by the CEO and CFO of the firm;
EX_OPTION = the number of unexercised options that the executive held at year-end that were vested scaled by total outstanding shares of the firm;
UN_OPTION = the number of unexercised options (excluding options grants in the current period) that the executive held at year-end that had not vested scaled by total outstanding shares of the firm;
GRNT_OPTION = new option grants made during the current period scaled by total outstanding shares of the firm; and
OWNER = the sum of restricted stock grants in the current period and the aggregate number of shares held by the executive at year-end (excluding stock options) scaled by total outstanding shares of the firm.

The t-statistics are calculated using Huber-White standard errors.

documented in Table 4, column 2. Finally, including interaction terms between the variables ΔGDP and *BIG* with the time variables did not affect the results reported in Tables 3 and 4.

Overall, the multivariate analysis can be summarized as follows: once we control for changes in the compensation structure, we do not find support for the claim that the SCA period was, on average, characterized by significantly higher accrual-based earnings management activities. However, we find higher real earnings management in the SCA period. These results suggest that there was an (as yet unidentified) exogenous force that affected the compensation structure and also either directly or indirectly (via the compensation structure) affected accrual-based earnings management.

After SOX, the level of accrual-based earnings management declined, but there was a significant increase in the overall level of real earnings management activities. Though the analysis does not allow inferences on whether the overall level of earnings management changed after SOX, it does indicate that firms seemed to switch to more real management activities after SOX, possibly since such methods are harder to detect. Finally, we find that compensation derived primarily from unexercisable stock options contributes significantly to the incentives of executives to manipulate earnings, a result consistent with recent research (e.g., Cheng and Warfield 2005).

SUSPECT Firm Analysis

One concern in the above analysis is whether our proxies are capturing earnings management activities of firms, or whether these trends that we observe represent some other phenomenon. To provide construct validity for our proxies, we conduct additional tests using "suspect firms" (SUSPECT) that are likely to have managed earnings based on three benchmarks that firms are likely to have incentives to meet. Using this subsample we examine whether earnings management strategies (i.e., the values of discretionary accruals, *DA*, and *RM_PROXY*) to meet these benchmarks have changed in the post-SOX period as compared to the pre-SOX period.

First, as in Roychowdhury (2006), we identify firm-year observations with net income before extraordinary items scaled by total assets that lies in the interval [0, 0.005) (we label these as SUSPECT firm-years), since it is likely that these firms during these years managed their earnings to report income marginally above zero. Table 5, Panel A reports the earnings management before and after SOX for firms that "just" managed to avoid reporting a loss.

Next, we use a second measure of SUSPECT firm-years when the change in net income before extraordinary items scaled by total assets lies in the interval [0, 0.005). This latter definition is consistent with evidence in prior research that firms are more likely to manage earnings in order to meet prior years' earnings numbers (Graham et. al. 2005). The results for this measure are reported in Table 5, Panel B.

Finally, recent research has suggested that meeting/beating analysts' forecasts is an important benchmark for firms and firms are likely to manipulate earnings to achieve this (e.g., Degeorge et al. 1999; Brown and Caylor 2005; Graham et al. 2005). Consistent with Roychowdhury (2006), we examine the accrual and real earnings management activities of firms that managed to meet or beat the final analysts' consensus forecast outstanding prior to the earnings announcement date (we use this consensus forecast as the *ex post* proxy for

TABLE 5

Earnings Management Activities of SUSPECT Firms in the Pre- and Post-SOX Periods

	Prior to SOX (1)	After SOX (2)	Difference (1)–(2)
Panel A: Accrual-Based and Real Earnings Management Firms that "Just" Avoided Reporting Losses			
DA	0.073	0.052	0.021***
RM_PROXY	−0.114	0.010	−0.214***
Panel B: Accrual-Based and Real Earnings Management Firms that Managed to Meet or Beat Last Year's Net Income			
DA	0.048	0.029	0.019***
RM_PROXY	−0.103	0.004	−0.107***
Panel C: Accrual-Based and Real Earnings Management Firms that Managed to Just Meet or Beat Analyst Forecasts by a Cent per Share			
DA	0.032	0.024	0.008**
RM_PROXY	−0.087	−0.034	−0.053***

*, **, *** Significant at 10 percent, 5 percent, and 1 percent levels, respectively.

Panel A: Firm-years "Just" avoiding reporting a loss are defined as firm-year observations where net income before extraordinary items scaled by total assets lies in the interval [0, 0.005).

Panel B: Firm-years that "Meet-or-Beat" last year's net income are defined as firm-year observations where the change in net income before extraordinary items scaled by total assets lies in the interval [0, 0.005).

Panel C: Firm-years that "Meet or Beat" analyst forecasts by one cent per share or less.

DA = discretionary accruals computed using the Modified Jones Model; and

RM_PROXY = sum of the standardized three real earnings management proxies, i.e., *R_CFO*, *R_PROD*, and *R_DISX*.

what managers expect the final consensus forecast to be for the year).[27] We obtain annual analyst forecasts from I/B/E/S and consider only the forecasts made and/or revised after the beginning of the fiscal year. We define the forecast error (*FE*) as the difference between actual earnings per share (EPS) as reported by I/B/E/S less the consensus forecast of earnings per share. We focus on firm-year observations where the *FE* is one cent per share or less ($\$0.00 \leq \text{EPS} - \text{Consensus forecast} \leq \$0.01$). We then define *SUSPECT_FE* as a binary variable taking the value of 1 if FE is one cent per share or less, and 0 otherwise. Table 5, Panel C presents the results of this analysis.

For each of the above three subgroups, we also compute earnings management for firms that just missed the three earnings targets and for all other firms that neither just missed nor just made the income targets. Because the suspect firms are likely to have engaged in income-increasing earnings management, we find that (not surprisingly) both before and after SOX, all three groups of suspect firms had significantly higher discretionary accruals when compared to either comparison group (results not tabulated). However, all firms that just avoided a loss (Panel A), that just managed to meet-or-beat last year's earnings (Panel B), as well as those that managed to meet or beat the consensus analyst forecast (Panel C) used significantly less income-increasing accrual-management after SOX when compared to firms in the same bin-intervals before SOX. An analysis of the real earnings management behavior of these firms indicates that the SUSPECT firms used significantly lower real earnings management than the other firms both before and after SOX (untabulated). However, real earnings management for these firms is significantly higher after SOX when compared to the firms in the same bin-intervals prior to SOX. This analysis provides us with greater confidence in the earlier results and strengthens our argument for the substitution between accrual and real earnings management methods after SOX.

V. SUMMARY AND CONCLUSION

We document both real and accrual-based changes in earnings management over time and examine whether the passage of SOX affected earnings management activities. Our results indicate that earnings management increased steadily over the sample period, and meeting or beating prior year's earnings numbers, consensus analysts' forecasts, and avoiding losses continued to be important incentives to manage earnings. We find an increase in earnings management in the period preceding the passage of SOX.

Our evidence suggests that subsequent to the passage of SOX, the level of earnings management returned to the pre-SOX trend line. We also find that while the level of accrual-based earnings management declined, the level of real earnings management activities increased significantly after the passage of SOX, suggesting that firms shifted from using accrual-based to real earnings management after SOX. Finally, our results indicate that the increase in accrual-based earnings management prior to SOX was concurrent with a shift in the fraction of equity-based executive compensation.

Our analysis does not permit us to attribute the above changes in earnings management activities solely to SOX, as several concurrent developments are likely to have contributed to a decrease in earnings management activities. For example, the Department of Justice began a series of highly publicized enforcement actions that are likely to have reduced managers' propensity to manage earnings. Similarly, the major governance failures are likely to have increased investors' and auditors' vigilance. Thus, although most of these other concurrent events are indirectly related to the regulations in SOX, we are cautious in making any causality inferences.

---

[27] Note that as an earnings benchmark, meeting/beating analyst forecasts are different from meeting/beating the zero earnings benchmark and loss avoidance. Given that analysts revise their forecasts several times in a year, this forms a moving target and it is not obvious as to which specific forecasts managers regard as the target they need to beat.

Finally, consistent with the results in Cheng and Warfield (2005) and Bergstresser and Philippon (2006), we interpret our evidence that the opportunistic behaviors of managers was one of the major antecedents of accrual-based earnings management during the period leading to SOX—primarily related to dramatic increases in the fraction of compensation derived from stock options. Consistent with the hypothesis that option components have a different set of incentives with regard to the exercise price, our evidence shows that while new options granted during the current period are negatively associated with income-increasing discretionary accruals, unexercised options are positively associated with income-increasing discretionary accruals. Furthermore, we document that the relation between income-increasing discretionary accruals and unexercised options (excluding new option grants) declined in the post-SOX period.

REFERENCES

Aboody, D., and R. Kasznik. 2000. CEO stock option award and the timing of corporate voluntary disclosures. *Journal of Accounting and Economics* 29: 73–1000.

Bartov, E., D. Givoly, and C. Hayn. 2002. The reward to meeting or beating analysts' forecasts. *Journal of Accounting and Economics* 33: 173–204.

Bergstresser, D., and T. Philippon. 2006. CEO incentives and earnings management. *Journal of Financial Economics* 80: 511–529.

Brown, L. D. 2001. A temporal analysis of earnings surprises: Profits versus losses. *Journal of Accounting Research* 39: 221–241.

———, and M. L. Caylor. 2005. A temporal analysis of quarterly earnings thresholds: Propensities and valuation consequences. *The Accounting Review* 80: 423–440.

Bumiller, E. 2002. Bush signs bill aimed at fraud in corporations. *N.Y. Times* (July 31).

Cheng, Q., and T. D. Warfield. 2005. Equity incentives and earnings management. *The Accounting Review* 80: 441–476.

Coffee, J. 2003. What causes Enron? A capsule social and economic history of the 1990s. Working paper, Columbia University.

Collins, D., and P. Hribar. 2002. Errors in estimating accruals: Implications for empirical research. *Journal of Accounting Research* 40: 105–134.

Dechow, P. M., R. G. Sloan, and A. P. Sweeney. 1995. Detecting earnings management. *The Accounting Review* 70: 193–225.

———, S. P. Kothari, and R. Watts. 1998. The relation between earnings and cash flows. *Journal of Accounting and Economics* 25: 133–168.

———, and D. J. Skinner. 2000. Earnings management: Reconciling the views of accounting academics, practitioners, and regulators. *Accounting Horizons* 14: 235–250.

———, and I. D. Dichev. 2002. The quality of accruals and earnings: The role of accrual estimation errors. *The Accounting Review* 77: 35–59.

DeFond, M. L., and J. Jiambalvo. 1994. Debt covenant effects and the manipulation of accruals. *Journal of Accounting and Economics* 17: 145–176.

Degeorge, F., J. Patel, and R. Zeckhauser. 1999. Earnings management to exceed thresholds. *Journal of Business* 72 (1): 1–33.

Erickson, M., M. Hanlon, and E. L. Maydew. 2006. Is there a link between executive equity incentives and accounting fraud? *Journal of Accounting Research* 44: 113–144.

Fields, T. D., T. Z. Lys, and L. Vincent. 2001. Empirical research on accounting choice. *Journal of Accounting and Economics* 31: 255–307.

Financial Accounting Standards Board (FASB). 1997. *Statement of Cash Flows*. Statement of Financial Accounting Standards No. 95. Norwalk, CT: FASB.

Forbes. 2002. The Corporate Scandal Sheet. Available at: http://www.forbes.com/home/2002/07/25/accountingtracker.html.

Fuller, J., and M. C. Jensen. 2002. Just say no to Wall Street. *Journal of Applied Corporate Finance* 14: 41–46.

Graham, J. R., C. R. Harvey, and S. Rajgopal. 2005. The economic implications of corporate financial reporting. *Journal of Accounting and Economics* 40: 3–73.

Greenspan, A. 2002. *Federal Reserve Board's Semiannual Monetary Policy Report to the Congress*. Testimony before the Committee on Banking, Housing, and Urban Affairs, U.S. Senate, July 16. Washington, D.C.: Government Printing Office.

Gunny, K. 2005. What are the consequences of real earnings management? Working paper, University of Colorado.

Healy, P. 1985. The effect of bonus schemes on accounting decisions. *Journal of Accounting and Economics* 7: 85–107.

———, and J. M. Whalen. 1999. A review of the earnings management literature and its implications for standard setting. *Accounting Horizons* 13: 365–384.

Jain, P. K., J. Kim, and Z. Rezaee. 2003. Have the Sarbanes-Oxley Act of 2002 and the CEO certifications made the market participants more informed? Working paper, University of Memphis.

———, and Z. Rezaee. 2006. The Sarbanes-Oxley Act of 2002 and capital-market behavior: Early evidence. *Contemporary Accounting Research* 23 (3): 629–654.

Johnson, S. A., H. E. Ryan, Jr., and Y. S. Tian. 2005. Executive compensation and corporate fraud. Working paper, Texas A&M University.

Jones, J. 1991. Earnings management during import relief investigations. *Journal of Accounting Research* 29: 193–228.

Kasznik, R. 1999. On the association between voluntary disclosure and earnings management. *Journal of Accounting Research* 33: 353–367.

Kothari, S. P., A. J. Leone, and C. Wasley. 2005. Performance matched discretionary accrual measures. *Journal of Accounting and Economics* 39: 163–197.

Lobo, G., and J. Zhou. 2006. Did conservatism in financial reporting increase after the Sarbanes-Oxley Act? Initial evidence. *Accounting Horizons* 20: 57–74.

Lopez, T. J., and L. L. Rees. 2001. The effect of meeting analyst forecasts and systematic positive forecast errors on the information content of unexpected earnings. Working paper, Georgia State University.

Matsumoto, D. 2002. Management incentives to avoid negative earnings surprises. *The Accounting Review* 77: 483–514.

Melancon, B. C. 2002. A new accounting culture. September 4. Available at: http://www.aicpa.org.

Paulson, H. 2007. The key test of accurate financial reporting is trust. *Financial Times* (May 17).

Rezaee, Z. 2004. Corporate governance role in financial reporting. *Research in Accounting Regulation* 17: 107–149.

Ribstein, L. E. 2002. Market vs. regulatory responses to corporate fraud: A critique of the Sarbanes-Oxley Act of 2002. *Journal of Corporation Law* 28: 1–68.

Roychowdhury, S. 2006. Earnings management through real activities manipulation. *Journal of Accounting and Economics* 42: 335–370.

Yermack, D. 1997. Good timing: CEO stock option award and company news announcements. *Journal of Finance* 52 (2): 449–476.

Zang, A. Z. 2006. Evidence on the tradeoff between real manipulation and accrual manipulation. Working paper, Duke University.

THE ACCOUNTING REVIEW
Vol. 83, No. 3
2008
pp. 789–822

Optimal Team Size and Monitoring in Organizations

Pierre Jinghong Liang
Carnegie Mellon University

Madhav V. Rajan
Stanford University

Korok Ray
University of Chicago

ABSTRACT: We formulate and analyze a model of team structure and monitoring within a Linear-Exponential-Normal (LEN) agency framework. We incorporate three key instruments in the internal design of an organization involving team production: team size, monitoring activities, and incentive contracts. We show that the complex trade-offs among these instruments lead to surprisingly simple implications. One such result is that the equilibrium level of pay-for-performance for workers is attenuated and is, at times, invariant to most environmental variables of interest. As such, our model helps explain the empirical puzzle of the lack of a trade-off for risk/incentives shown in standard agency models. Our work also demonstrates the presence of complementarities between team size and monitoring, and between worker talent and managerial monitoring ability. Finally, we derive predictions about the impact of environmental variables on the choice of optimal team size, incentives, and employee quality, even in the presence of an external marketplace for talent.

Keywords: *incentives; team size; monitoring.*

I. INTRODUCTION

Teams play a pivotal role in modern organizations. The study of teams has occupied a similarly important place in modern economic thought, beginning with the work of Marschak and Radner (1972). The primary focus of inquiry has been the provision of incentives to motivate team members to overcome concerns regarding "free-riding." Virtually all of this work, following the lead of Holmström (1982), takes team structure and size as given. Yet, the broader elements of organizational design also form a key component of modern firms (see Roberts 2004). Management consultants and the business

---

We are grateful to Jon Glover, Steve Huddart, Stefan Reichelstein, Amir Ziv, and three anonymous reviewers for many helpful suggestions. We have also benefited from comments by Jeremy Bertomeu, Shane Dikolli, Sunil Dutta, Qintao Fan, Nicole Johnson, Dhananjay Nanda, Richard Saouma, Haresh Sapra, Jason Schloetzer, and seminar participants at University of California, Berkeley, Carnegie Mellon, University of Chicago, City University of Hong Kong, Duke, and the Issues in Managerial Accounting Conference at The University of Texas at Austin.

Editor's note: This paper was accepted by Steven Huddart.

Submitted March 2006
Accepted October 2007

press increasingly exhort executives to view organizational design as yet another choice variable. The academic press has only recently begun to explore such notions (Simons 2005), and theoretical research on the topic is still nascent (see, e.g., Harris and Raviv 2002). Acemoglu et al. (2006) note that the growing interest in the determinants of firms' organizational choices is partly motivated by the belief that recent technological developments are favoring more decentralized firms.

In this paper, we provide a model of team formation and monitoring within the Linear-Exponential-Normal (LEN) agency framework that enables us to shed light on optimal incentives, team size, and the efficient matching of worker and managerial abilities. We consider three instruments in the internal design of an organization involving team production: team size, monitoring activities, and the incentive contracts offered to workers and managers. The complex interactions among the three instruments generate new insights into the optimal provision of incentives. One such insight is that when the error in the measurement of employee contribution to firm value satisfies mild regularity conditions, incentives are attenuated and in some cases optimally structured to be invariant to noise, risk preferences, and the cost of worker effort. Our findings offer a novel explanation for the lack of empirical link between risk and incentives (see, e.g., Bhattacharyya and Lafontaine 1995; Prendergast 2002).

A central feature of modern organizations is the careful design and implementation of internal and external monitoring activities. These activities influence (and are influenced by) other organizational design choices, such as team size and incentive schemes. In this study, we examine the optimal number of risk-averse workers and the nature of their incentives when the principal hires a risk-averse manager to watch over the workers. Much of the previous literature on monitoring has assumed that monitoring is either exogenous or is done by the principal herself. In this paper, we deviate from this (unrealistic) assumption and, by using a LEN formulation, generate closed-form solutions to the incentive problems among the principal, the self-interested manager, and the workers being monitored. We derive the determinants of optimal team size and characterize the impact of this choice on the strength of the incentives given to managers and workers. We also identify conditions under which it is optimal to employ a manager, and demonstrate the existence of endogenous complementarities across the skill levels of the manager and workers. Finally, we provide comparative statics results on the impact of changes in uncertainty and performance measurement technology on the efficient mix of worker and manager talent employed by the firm.

Given our choice to emphasize three main instruments—team size, monitoring, and incentives—in our analysis, our results are more likely to be relevant to firms with organizational features that emphasize these three dimensions. We believe modern knowledge-based industries fit our broad description well. They are more likely to be organized into teams of varying sizes; their incentives are more likely to be output based; and professional managers are engaged in monitoring efforts from individual workers. As an example, consider a software manager who supervises a team of software engineers. As the team grows, it becomes more difficult to coordinate these multiple engineers and this makes performance measurement harder. Our paper speaks to the efficient provision of incentives for the supervisor and employees, the limits placed on the size of teams by incentive and monitoring considerations, and the optimal mix of talent levels across workers and managers.

Some authors, notably Williamson (1967), have suggested that the inherent loss of control from hiring monitors ultimately limits the size of the firm. Calvo and Wellisz (1978) formalize this argument and demonstrate that Williamson's results require particular *ad hoc* assumptions on the firm's monitoring technology. For example, if agents are unaware of when they are being monitored, then Calvo and Wellisz (1978) find that there is no hierarchical limit to firm size. In our model, the active monitoring reduces the variance of the group performance measure. As such, our managers can be interpreted as accountants or auditors whose job is to reduce measurement errors in performance metrics. The reduction in noise allows the principal to offer stronger incentives to the workers without the burden of paying additional risk premia. In our model, there is no loss of control with the delegation of monitoring to a third party, as our entire model is cast in a moral hazard setting.

Because our monitors are separate from the principal, our framework differs from the traditional role assigned to monitors in the accounting literature (see, e.g., Baiman and Demski 1980). More recent work in the economics literature has considered the possibility of having working agents monitor one another via either direct observation (Ma 1988; Varian 1990) or via indirect observation of correlated signals (Fischer and Hughes 1997). In his paper, Varian (1990) shows that it may be optimal to accept only groups of agents, because the agents will only select "good types" for their own group. In Section III, we allow for heterogeneous types of workers and managers and find that as risk increases, rather than selecting a better monitor, the principal will want to decrease the team size, and hire a lower quality monitor.

We utilize a multi-agent LEN framework to analyze issues related to optimal team incentives. Perhaps the earliest paper to employ this approach was McLaughlin (1994), who looked at the determinants of the links between individual compensation and team performance. In Adams (2006), the productive complementarities among team members in team production make the use of firm-wide incentives valuable, even in the presence of the induced free-rider problem. Recent work has extended the analysis of team incentives to incorporate elements related to optimal task assignments (Hughes et al. 2005) and the role of productive synergies in teams (Autrey 2005). Hughes et al. (2005) find sufficient conditions for an owner to prefer diverse task assignments as opposed to specific assignments. Autrey (2005) demonstrates that an owner favors the measurement of individual output when there is little synergy across workers' efforts, but prefers to compensate employees based on team output once the level of synergy is sufficiently high. None of the above papers incorporates managerial monitoring or allows for the endogenous determination of team size and composition.

Qian (1994) is one of the few studies that allows for an endogenous number of workers. Following Williamson (1967) and Calvo and Wellisz (1978), Qian examines the optimal size of a hierarchy and compensation throughout all tiers of the hierarchy. In Qian's (1994) paper, the principal benefits from having fewer tiers in the hierarchy because it is assumed that there are production losses at each tier and, in general, fewer tiers imply fewer managers to compensate. On the other hand, if the firm has fewer tiers, then each manager's span of control is enlarged and hence monitoring becomes less effective; the principal is thus forced to pay a higher efficiency wage to ensure that the managers put forth costly monitoring efforts. In our paper, the manager serves to reduce the noise in measuring workers' effort, and hence we only consider a two-tier hierarchy: principal at the top, and workers accompanied by the manager at the bottom. In Qian's (1994) model, the agents are assumed to be risk-neutral with limited liability and are compensated with two-tier wages. By contrast, we assume that all workers and the manager are risk-averse, and are compensated via the optimal linear contract.

From the standpoint of the accounting literature, our work is most closely related to Ziv (1993, 2000). Ziv (1993) considers the interaction between a firm's internal information structure and its optimal size. The study focuses on how the induced effort level and firm size respond to alternative, exogenous, information structures (such as whether signal variance is constant or increasing in firm size). Ziv (2000) introduces self-interested managers and studies the optimal number of layers of a firm's hierarchy and the number of agents in each layer. In his model, the principal can monitor the agents herself, which eliminates the incentive problem with monitoring. Hiring supervisors to monitor the lower-layer agents, on the other hand, introduces such an incentive problem. Thus, Ziv (2000) concludes that "flatter" organizations are optimal in most cases.[1]

In our paper, we assume that the absentee principal is unable to engage in monitoring directly, and must hire one or more self-interested managers and motivate them to engage in monitoring. Comparing this setting with a benchmark (called mechanical monitoring) allows us to clearly establish the trade-offs inherent in hiring a manager (i.e., the gains from improving the worker's incentive versus the incentive problem associated with the manager's monitoring). In addition, we are able to characterize incentive contracts for the workers and manager, and can determine, in closed-form, the optimal number of workers. In turn, this allows us to demonstrate the presence of organizational synergies when manager and worker abilities are exogenously given, as well as in cases where the firm must hire employees of varying abilities (and associated reservation wages) in a marketplace for talent.

Huddart and Liang (2005) study related issues in the context of a partnership. In a partnership, the sharing of profits invites all partners to shirk their productive efforts. While monitoring activities may improve the incentive situation, they may also lead to additional shirking because the partners also act as monitors. A key difference in our paper is the presence of a risk-neutral principal who eliminates the risk-sharing considerations inherent in a partnership setting. In addition, we model managers as agents specialized in monitoring, while not directly involved in actual production, thereby capturing real-world settings beyond those in professional service firms organized as partnerships. Our setting allows us to analyze the nature of the interaction among design instruments (incentive contracts, worker size, and monitoring assignments), how such interaction leads to the attenuation of worker incentives, and the value of hiring a manager.

The remainder of the paper is organized as follows. Section II lays out the basic model and characterizes its solution under both the benchmark (mechanical) and managerial monitoring settings. Section III derives the existence of complementarities in the selection of managerial and worker talent, in particular when external labor markets are present, and provides additional comparative statics results. The last section offers concluding comments and suggests ways in which the analysis in this paper could be extended or generalized.

II. THE MODEL

Consider a principal (the firm) contracting with n workers. Each worker exerts effort e_i at cost of effort $C_i(e_i) = \frac{1}{2}c_i e_i^2$. The cost of effort parameter c_i measures the quality/ability of the worker: higher quality workers have a lower c_i. Let $N = \{1, 2, \ldots, n\}$ denote the set of workers. Joint output is given by:

$$x = \sum_{i \in N} e_i + \varepsilon_x.$$

---

[1] Also see Baldenius et al. (2002) for related work along these lines.

We assume that x is either unverifiable or is realized too late to be used in contracting. Instead, the principal can install a mechanical monitoring system at a price (normalized to zero), which produces a signal y.[2] This is the only variable on which contracts can be written. We specify that:

$$y = \sum_{i \in N} e_i + \varepsilon, \text{ where } \varepsilon \sim N(0, G(n)\sigma^2).$$

Note that σ measures the uncertainty in performance measurement. We normalize $G(0)$ to equal 0, and assume that $G(n)$ is increasing and strictly convex in n, so that measurement becomes noisier as the team size grows. This reflects a basic assumption on the monitoring technology, namely that the performance of larger teams is harder to measure because of problems with coordination, communication difficulties, etc. In addition, we assume that $G(\cdot)$ satisfies the following technical condition:

$$\frac{d}{dn}\left(\frac{n \cdot G'(n)}{G(n)}\right) \geq 0. \tag{1}$$

This condition is not too restrictive, and is satisfied by a wide range of functional forms. It holds for all power functions ($G(n) = An^{\gamma}$, $A > 0$, $\gamma > 1$), as well as any general polynomial function with positive coefficients. It is also satisfied by functions in the exponential class ($G(n) = A(e^n - 1)$, $A > 0$). Finally, note that the condition is implied by, and is therefore weaker than, a requirement that $G(n)$ be weakly log-convex.

The firm offers a linear contract to each worker, so each worker receives a wage $w_i = a_i + b_i y$. All n workers are equally risk-averse with exponential utility and a common coefficient of risk aversion parameter r. Thus, workers' preferences assume a mean-variance representation. In certainty equivalent terms, each worker receives $\mathrm{E}w_i - \frac{r}{2}\mathrm{Var}(w_i) - C_i(e_i)$: expected wages less a risk premium less the cost of effort. Each worker solves:

$$\max_{e_i} a_i + b_i \mathrm{E}y - \frac{r}{2} b_i^2 G(n)\sigma^2 - C_i(e_i),$$

yielding the incentive constraint, $e_i = b_i/c_i$.[3] Suppose each worker has an outside option, denoted by $\bar{u}$, which is normalized to zero.[4] The firm will set the salaries (a_i) such that each worker's individual rationality constraint binds, given that all other workers select their equilibrium effort levels.

Assume each unit of output (x) is sold in a competitive market at an exogenous price q. The firm maximizes the expectation of revenues less wage payments, and substitutes the binding (IR) constraints into its optimization. So the firm selects bonus coefficients (incentives) to maximize total surplus:

---

[2] We present this benchmark case in order to highlight the costs and benefits that emerge when a self-interested manager is hired to provide the informative signal about workers' efforts.

[3] Notice that for each worker, the effort choice is also a dominant strategy response in the workers' subgame.

[4] We relax this assumption in Section III.

$$\max_{b_i} \sum_{i \in N} \left[qe_i - C_i(e_i) - \frac{r}{2} b_i^2 G(n)\sigma^2 \right].$$

This objective function reveals the two costs of production that the principal must pay the workers. The first is the personal cost of providing effort, $C_i(e_i)$. The second is the risk-premium, $\frac{r}{2} b_i^2 G(n)\sigma^2$. This is the incentive cost, which is the main focus of our study. After inserting the worker's incentive constraint, the first-order condition gives the optimal incentives for an exogenous n:

$$\hat{b}_i(n) = \frac{q}{1 + rc_i G(n)\sigma^2}, \tag{2}$$

yielding the standard risk/incentive trade-offs. Observe that a larger team leads to a higher variance of the team performance measure, which leads to the firm cutting back on incentives. Also note that the equilibrium effort from each worker is $\hat{b}_i / c_i$, which is lower than the first-best effort level of q / c_i.

Incentives and Optimal Team Size

In the standard LEN model (e.g., a single-agent, single-task, single-period model), the optimal incentive weight is $\hat{b} = 1/(1 + rc\sigma^2)$ and the pay-for-performance sensitivity $\partial\hat{b}/\partial\sigma$ captures how incentives vary with the risk or uncertainty in the model. In the standard model this derivative is negative, suggesting a trade-off between risk and incentives. Yet the empirical support for this trade-off is weak, as noted in the introduction. We show now that adjusting for team size attenuates the effect of risk/uncertainty on incentives, which may help explain why a strong negative trade-off is not observed empirically.

In our model, assuming identical workers ($c_i = c$) leads to $\hat{b}_i(n) = \hat{b}(n)$, which in turn allows us to solve for the optimal incentives and total profit for a given n:

$$\hat{b}(n) = \frac{q}{1 + rcG(n)\sigma^2} \quad \text{and} \quad \hat{\Pi}(n) = \frac{nq\hat{b}(n)}{2c}. \tag{3}$$

The firm chooses an optimal team size, denoted $\hat{n}$, which maximizes $\hat{\Pi}(n)$. As $\hat{\Pi}$ and its derivative are continuous on $[0, \infty)$ and $\hat{\Pi}'(0) > 0$, a necessary condition for $\hat{n}$ to be an optimal team size is that $\hat{\Pi}'(\hat{n}) = 0$. This yields the following first-order condition:

$$\hat{n}G'(\hat{n}) - G(\hat{n}) = \frac{1}{rc\sigma^2}. \tag{4}$$

Since the expression on the left is 0 at $n = 0$ and is strictly increasing in n (from the convexity of $G(\cdot)$), a solution, $\hat{n}$, exists and is unique. Furthermore, it is easy to verify that $\hat{\Pi}''(\hat{n}) < 0$, so that $\hat{n}$ is indeed the optimal team size.

Implicit differentiation in Equation (4) yields our first key result:

Proposition 1: The optimal team size decreases as either worker risk aversion r, worker cost of effort c, or uncertainty in performance measurement σ^2 rises.

The ability to adjust team size optimally in turn alters the optimal adjustments in incentives. We now examine how incentive strength $\hat{b}$ reacts to changes in uncertainty σ^2, worker risk aversion r, and worker cost of effort c when team size is adjusted optimally, compared to the case when team size is exogenous. Let us denote any of the symbols r, c, or σ^2 by v. Suppose the current team size is n_0. If the team size were determined exogenously, then the effect of changes in v would be given by:

$$\delta^{v}_{\text{ex}}(n_0) \equiv \left.\frac{\partial \hat{b}}{\partial v}\right|_{n=n_C}, \tag{5}$$

where $\hat{b}$ is given by Equation (3). If team size is chosen optimally, the effect of changes in v is given by:

$$\delta^{v}_{\text{end}}(n_0) \equiv \left.\frac{\partial \hat{b}}{\partial v}\right|_{n=n_0} + \left.\frac{\partial \hat{b}}{\partial n}\right|_{n=n_0} \cdot \left.\frac{\partial \hat{n}}{\partial v}\right|_{\hat{n}=n_0}, \tag{6}$$

where $\partial\hat{n}/\partial v$ is obtained by implicit differentiation in Equation (4). We show next that when team size is adjusted efficiently in reaction to changes in external factors, the directional impact on the incentive strength offered to workers is unaltered, but the magnitude of the impact is dampened.

Proposition 2: For any set of parameter values (σ^2, r, c), $\delta^{v}_{\text{ex}}(\hat{n}) < \delta^{v}_{\text{end}}(\hat{n}) \leq 0$.

Stated differently, $|\delta^{v}_{\text{end}}(\hat{n})| < |\delta^{v}_{\text{ex}}(\hat{n})|$, i.e., pay-for-performance sensitivity under optimally chosen team size is *always* lower than it would be if the team size were given exogenously. In contrast to the standard agency model, the firm here has two instruments: team size and incentives. An increase in r, c, or σ^2 results in a direct decrease in $\hat{b}$ and an indirect increase in $\hat{b}$ through the decrease in $\hat{n}$. Therefore, the direct effect of a decrease in incentives in response to a change in environmental parameters is partially offset by an increase in incentives due to the indirect effect of a decrease in team size (incentives being decreasing in team size.)

An important special case of the attenuation of incentives occurs when $G(\cdot)$ is a power function, $G(n) = An^{\gamma}$ (with $A > 0$ and $\gamma > 1$). In this case the pay-for-performance sensitivity is not just attenuated, but eliminated entirely.

Proposition 3: The optimal incentive weight b does not depend on parameter values (σ^2, r, c) if and only if $G(\cdot)$ is a power function.

That is, incentives exhibit an *invariance property* when $G(\cdot)$ is a power function: the optimal incentives do not vary at all with σ^2, r, or c. The firm responds to changes in these parameters by adjusting the team size and keeping the incentives fixed.

To understand the intuition behind this result, suppose first that n is exogenous, and σ^2 increases. From Equation (3), $\hat{b}$ decreases, since the firm cuts incentives as the risk premium rises. As the incentives fall, so do the profits. Now if the firm decreases the team size, a smaller n reduces the risk premium and hence lifts the incentives. This rise in incentives offsets the fall in profits from the smaller team size. When $G(\cdot)$ is a power function, the firm decreases the team size *exactly* to offset the increase in uncertainty, leaving the incentives constant. Essentially, team size is a more flexible instrument for the firm than are incentives.

If $G(\cdot)$ is a power function, then the optimal team size can be derived in closed-form as:

$$\hat{n} = \left(\frac{1}{(\gamma - 1)Arc\sigma^2}\right)^{1/\gamma}. \tag{7}$$

Substituting this expression into the equation for the workers' incentives yields the equilibrium strength of incentives and total profit at the optimal team size:

$$\hat{b}(\hat{n}) = q \cdot \left(\frac{\gamma - 1}{\gamma}\right) \text{ and } \hat{\Pi}(\hat{n}) = \frac{\hat{n}q^2(\gamma - 1)}{2c\gamma}. \tag{8}$$

From Equation (8), the invariance property is readily observable: the optimal incentive weight, $\hat{b}(\hat{n})$, is independent of σ, r, and c.

Before ending our discussion of the benchmark model, we comment on the robustness of the above findings. Two aspects of our formulation invite special examination—the use of a single group measure to compensate each worker, and the specification of an additively separable production function. Similar results (e.g., the attenuation and invariance findings, and the impact of external factors on team size) continue to hold when each of these assumptions is altered. In particular, the results generalize immediately to a setting where each worker is compensated using his own *individual performance signal*, as opposed to the common group measure y. They also extend to a scenario where the production technology exhibits complementarities among workers, as opposed to the linear technology with uniform marginal productivity (q) in the benchmark model.[5]

Adding a Manager

Now we introduce a manager to the benchmark setting. The manager serves the role of actively and endogenously supplying the monitoring activities (as opposed to a mechanical monitoring system). Endogenous monitoring alters the economic trade-off on team size and worker incentives (and thus welfare), but we show that the attenuation result continues to hold when team size is endogenously chosen.

We model the manager as an effort- and risk-averse economic agent who can produce a signal about the collective efforts of the team under his supervision. In particular, the manager exerts monitoring effort m at personal cost $C(m) = k \cdot m$, where k is the manager's cost of effort. The parameter k measures the manager's quality/ability: higher quality (or lower cost) managers have a lower k. The manager's effort reduces the variance on the team performance measure. So, while we use the term manager, we could equivalently refer to a monitor or an internal auditor, as the function of such an individual is to improve the quality of the measurement of performance of other workers within the firm. Specifically, we assume:

---

[5] Details of these results are available from the authors on request. More generally, our particular modeling choice (i.e., placing the team size effect on the variance of the performance measure) is, in fact, fairly general. One can show that placing the team size effect on worker-effort costs, performance measure sensitivity, or the marginal benefit of effort are all equivalent ways of representing the model. We are grateful to one of our reviewers for alerting us to this point on scaling, and for demonstrating it mathematically.

$$y = \sum_{i \in N} e_i + \varepsilon, \text{ where } \varepsilon \sim N\left(0, \frac{G(n)\sigma^2}{m}\right).$$

Task-specialization is such that the manager's effort reduces the variance of y while the worker's effort increases its mean.[6] Assume m is unobservable so the manager must be motivated to supply monitoring efforts. The firm offers a linear contract to the manager consisting of a salary and a bonus on the team performance measure,[7] i.e., $\omega = \alpha + \beta y$. The manager thus solves:

$$\max_{m} \alpha + \beta \mathrm{E}y - \frac{r}{2}\beta^2 \frac{G(n)\sigma^2}{m} - C(m).$$

Even though all workers and the manager are under linear contracts, the provision of incentives to exert effort comes through in different ways. The manager's effort does not affect the expectation of his wage, but such effort does decrease its variance. Given the manager is risk-averse, he has an incentive to reduce the variance. The first-order condition yields the incentive constraint for the manager:

$$m^* = \beta\sigma\left(\frac{G(n)r}{2k}\right)^{1/2}. \tag{9}$$

The manager's problem is concave even when his personal cost $C(m)$ is linear because his personal benefit of monitoring effort is increasing and concave in m.[8]

From Equation (9), we learn that, for a given incentive coefficient β, if the measurement process worsens (σ^2 increases) or the manager becomes more risk-averse (r increases), the risk premium grows and hence the manager works harder at monitoring. Moreover, as the manager's span of control grows (n increases) the manager also works harder. Adding workers boosts the variance of y and hence increases the risk premium. Because the manager dislikes variance, he spends more time monitoring to compensate for the additional risk cost. However, the manager's incentive coefficient β is endogenous, and depends on the parameters σ and r. We return to address the overall effect on equilibrium monitoring efforts later (in Proposition 5).

As before, the firm sets the salary levels (a_i, α) such that the individual rationality constraint binds for every worker and the manager. Assume the opportunity wage for the manager is $\bar{u}_M$, which is also normalized to zero. The firm therefore maximizes total surplus, as before:

---

[6] By this modeling choice, we choose to emphasize the monitoring role of managers, leaving their other functions (such as fulfilling direct productive tasks) as second-order effects. Within this monitoring role, the manager can, for example, exert effort to coordinate multiple agents within the team or facilitate communication across agents. As a reduced form, we model this as decreasing the variance on the team performance measure, as opposed to explicitly framing the exact consequence of the managerial action, in order to keep the level of complexity manageable.

[7] We believe this is a simple and natural way to model the compensation arrangement between the principal and the manager. In practice, managers' compensation is, in large part, determined by the performance of the group under their supervision; see, for example, Bushman et al. (1995). Providing managerial incentives via the signal produced by the manager represents a sharp contrast to the model in Ziv (2000), where the firm provides incentives to managers via a signal about managerial effort that is generated by yet another manager or, ultimately, the principal. While Ziv's (2000) model potentially allows for the investigation of multi-layer hierarchies, the imposed complexity precludes development of a tractable framework.

[8] A benefit of linear personal cost is that it provides tractability. Other functional forms of $C(m)$ provide similar economic intuition, but result in mathematical expressions of far greater complexity.

$$\max_{b_i,\beta} \sum_{i\in N}\left[qe_i - C_i(e_i) - \frac{r}{2}\mathrm{Var}(w_i)\right] - C(m) - \frac{r}{2}\mathrm{Var}(\omega),$$

where:

$$\mathrm{Var}(w_i) = b_i^2\,\frac{G(n)\sigma^2}{m} \qquad \text{and} \qquad \mathrm{Var}(\omega) = \beta^2\,\frac{G(n)\sigma^2}{m}.$$

Substituting in each worker's incentive constraint gives:

$$\max_{\beta,b_i} \sum_{i\in N}\left[\frac{qb_i}{c_i} - \frac{b_i^2}{2c_i} - \frac{r}{2}\,b_i^2\,\frac{G(n)\sigma^2}{m^*}\right] - km^* - \frac{r}{2}\,\beta^2\,\frac{G(n)\sigma^2}{m^*}, \tag{10}$$

where m^* is given by Equation (9). The first-order conditions yield the optimal incentives for the worker:

$$b_i^*(n) = \frac{q}{1 + \dfrac{rc_iG(n)\sigma^2}{m^*}}.$$

Similar to the earlier setting with a mechanical monitor, an increase in risk aversion (r), cost of effort (c_i), or noise (σ^2) causes the firm to cut the worker's incentives, delivering the standard trade-off. What is different now is that the manager's effort (m) matters in choosing optimal incentives for workers. Of course, the firm affects m by its choice of the manager's incentive β.

Optimizing the firm's objective function with respect to β results in a first-order condition, from which we can see the principal's trade-off in choosing the manager's incentive:

$$\left[\frac{r}{2}\,\frac{G(n)\sigma^2}{m^{*2}}\sum_{i\in N} b_i^2 - k + \frac{r}{2}\,\frac{G(n)\sigma^2}{m^{*2}}\,\beta^2\right]\frac{\partial}{\partial\beta}m^* - r\beta\,\frac{G(n)\sigma^2}{m^*} = 0. \tag{11}$$

Providing more managerial incentive (i.e., increasing β) induces more managerial effort (see Equation (9)). The benefit is two-fold. First, it reduces the variance of the performance measure y and thus the risk-premium needed for the manager because the manager's contract is written on y. Second, increasing β reduces the risk-premium of all workers because the worker contracts are also written on y. Notice here that the firm exploits a spill-over effect from the managerial monitoring effort. The manager is self-interested and wants to exert effort to reduce the variation in his own compensation. But in doing so, he reduces the wage variances of all workers. This is beneficial to the firm because it reduces salary costs (a_i^* and α^*). These two (marginal) benefits are shown as the two positive terms in Equation (11). The cost of providing managerial incentives is also two-fold. First, a higher incentive leads to more personally costly managerial effort, for which the manager requires compensation. Second, a higher incentive leads to managerial pay that is more sensitive to the signal variations, which increases the risk-premium. These two (marginal) costs are shown as the two negative terms in Equation (11).[9]

Substituting the manager's incentive constraint (9) into Equation (11), and rearranging terms yields:

$$\beta = \left(\frac{1}{2}\sum_{i\in N} b_i^2\right)^{1/2}.$$

The manager's incentives thus increase with each worker's incentives. Intuitively, as the firm increases b_i, it induces more effort from the workers but also increases the risk premium. To compensate, the firm increases β, which induces the manager to spend more effort monitoring the workers; in turn, this reduces the risk premium for everyone.

Combining the previous expressions for b_i, β, and the manager's incentive constraint yields the following expression for b_i^* as an implicit function of the model parameters:

$$b_i^*(n) = \frac{q}{1 + 2c_i\sqrt{\dfrac{rG(n)\sigma^2 k}{\sum_{j\in N} b_j^{*2}}}}. \tag{12}$$

Despite the fact that Equation (12) is an implicit function of b_i, it can be shown, as in the setting with a mechanical monitor, that the strength of incentives (b_i) falls as the cost of effort (c_i) increases.

Optimal Incentives and Team Size with a Manager

Assuming identical workers ($c_i = c$) leads to $b_i^* = b^*$, which in turn allows us to solve Equation (12) and derive a closed-form solution for the equilibrium worker incentive for a given n:

$$b^*(n) = q - 2c\sigma\sqrt{rkG(n)/n}. \tag{13}$$

Recall that the equilibrium manager incentive (β^*), monitoring effort (m^*), and worker effort (e^*) are all functions of b^*.[10] From Equation (10), we can now write the profit function with a manager for a given n, denoted by $\Pi(n)$, as follows:

---

[9] Consider a slight alteration of the model where the manager exerts a productive effort, say e_m, which, similar to the workers' efforts e_i, increases the output and the aggregate performance measure y. When choosing the optimal β, the principal now must consider the marginal benefit and cost of inducing e_m, in addition to the four effects consider in Equation (11). However, so long as these additional marginal effects are well behaved (i.e., smooth and bounded), their only impact is to make the optimal β quantitatively different. The fundamental trade-off in the managerial monitoring activities, which is the focus of our paper, persists and leads to the same qualitative results.

[10] The maintained assumption here is that $b^* > 0$. We rule out combinations of parameters such that the above equation gives a negative b^*, because that implies equilibrium worker incentives would be set to zero, effectively shutting down production.

$$\Pi(n) = n\left[\frac{qb^*}{c} - \frac{b^{*2}}{2c} - \frac{r}{2}b^{*2}\frac{G(n)\sigma^2}{\frac{b^*\sigma}{2}\sqrt{\frac{rnG(n)}{k}}}\right] - 2k\frac{b^*\sigma}{2}\sqrt{\frac{rnG(n)}{k}}$$

$$= \frac{n(b^*(n))^2}{2c}. \qquad (14)$$

Now we characterize the optimal n in this setting.

Lemma 1: Assume $c_i = c$ for all $i \in N$, and the principal hires a manager with ability k. There exists a unique optimal team size, n^*, which is characterized by the following implicit function:

$$h(n^*) + 2n^*h'(n^*) = \frac{q}{2c\sigma\sqrt{rk}},$$

where $h(n) = \sqrt{G(n)/n}$.

Implicit differentiation of the equation defining n^* above enables us to study how the firm adjusts team size in response to changes in the exogenous parameters of interest.

Proposition 4: The optimal team size is decreasing in measurement uncertainty σ, worker risk aversion r, and worker and manager costs of effort, c and k, respectively.

When either c, k, r, or σ increases, this inflates the risk premium. As in the mechanical monitor setting, the firm responds by choosing a smaller team, thus deflating the risk premium somewhat. As the quality of the manager improves (k shrinks), the optimal team size grows, for two reasons. The first is the reverse of the risk-premium effect mentioned above. Second, better managers are better able to measure performance, and can thus handle larger teams.

We next study whether the ability to adjust team size affects the strength of incentives, as in the mechanical monitoring setting. Let v denote any of the parameters σ, c, $\sqrt{r}$, or $\sqrt{k}$. The impact on the workers' incentive parameter, b, of factor v is then given b:

$$\varphi^v_{ex}(n_0) \equiv \left.\frac{\partial b^*}{\partial v}\right|_{n=n_0}$$

and:

$$\varphi^v_{end}(n_0) \equiv \left.\frac{\partial b^*}{\partial v}\right|_{n=n_0} + \left.\frac{\partial b^*}{\partial n}\right|_{n=n_0} \cdot \left.\frac{\partial n^*}{\partial v}\right|_{n^*=n_0},$$

for the exogenous and the endogenous team-size cases, respectively. Since the monitoring is now performed by a self-interested manager, an immediate question is whether a similar comparison can be performed on the strength of the incentive, β, offered to the manager.

We refer to the corresponding constructs for the manager as $\psi^{v}_{ex}(n_0)$ and $\psi^{v}_{end}(n_0)$, respectively. Finally, it is also of interest to identify the relative impact of parameter changes on the manager's and workers' equilibrium levels of pay-for-performance. The following result provides a full characterization of these issues.

Proposition 5: For any set of parameter values (σ^2, c, r, k), and for a change in any factor $v = \sigma$, c, $\sqrt{r}$, or $\sqrt{k}$:

$$\psi^{v}_{ex}(n^*) = \psi^{v}_{end}(n^*) \leq \varphi^{v}_{ex}(n^*) < \varphi^{v}_{end}(n^*) \leq 0,$$

with a strict inequality in the middle for any $n^* > 2$.

There are several implications embedded in the above string of inequalities. First, as in the mechanical monitor setting, $|\varphi^{v}_{end}(n^*)| < |\varphi^{v}_{ex}(n^*)|$, i.e., the introduction of team size as a choice variable unambiguously attenuates worker incentives. An increase in σ, c, r, or k results in a decrease in the strength of incentives offered to the workers, regardless of whether team size is exogenously given or optimally adjusted. In the latter case, however, changes in any of these exogenous factors leads to a smaller change in the b offered to workers, i.e., the impact of variations in the external environment on the optimal contract offered to workers is dampened.

Second, we find, rather surprisingly, that no such attenuation occurs for the manager's incentives. Starting from the efficient team size, the variation in the manager's incentive rate when any external factor is changed is identical regardless of whether team size is adjusted to the new optimal level. Moreover, the decrease in the manager's incentive coefficient for an increase in any factor, v, is greater than the corresponding drop in the workers' incentive coefficient. We thus find that, even though both workers and managers are risk-averse, the manager's incentives are unambiguously more sensitive to changes in exogenous parameters (uncertainty, risk aversion, or costs of effort) than are the incentives of the workers. Our model thus yields the testable prediction that one should empirically observe greater variation in the incentives for supervisors than those for production workers.

The intuition for the result is as follows. We know from the previous section that $\beta^* = b^*\sqrt{n/2}$. Thus, parameters affect the manager's incentives via two channels: through their effects on b^* and through their effects on the team size. The effect of the parameters on β^* is stronger than that on b^* whenever the effects on b^* and n work in the same direction. Proposition 4 shows that an increase in any of the parameters σ, c, r, or k causes the firm to shrink the size of the team. Thus, whenever the firm reduces b^* in response to these parameter changes, the effects of the parameters on b^* and n amplify each other; as a result, β^* responds more than does b^*.

We next show that the extreme version of attenuation, i.e., the invariance property for workers' incentives, continues to hold for power functions even in the presence of a manager.

Proposition 6: The optimal b does not depend on parameter values (σ, c, r, or k) if and only if G is a power function, $G(n) = An^{\gamma}$.

A different way of stating this result is as follows: the weak inequality, $\varphi^{v}_{end}(n^*) \leq 0$, in the statement of Proposition 5 is strict unless $G(n)$ belongs to the class of power functions. At an intuitive level, we know from Proposition 3 that the invariance property holds in this new setting if and only if $G(n)/m^*$ is a power function. But, by Equation (9), $m^*(n)$

is a power function if and only if $G(n)$ is a power function. Consequently, $G(n)/m^*$ is a power function if and only if $G(n)$ is.

Propositions 5 and 6 together imply that if $G(\cdot)$ is a power function, then the manager's incentives respond to parameter changes, while the incentives of the workers do not. To see this explicitly, consider the following closed-form solutions for the equilibrium levels of the workers' and manager's incentives when team size is chosen optimally:

$$b^*(n^*) = q \cdot \left(\frac{\gamma - 1}{\gamma}\right) \tag{15}$$

and:

$$\beta^*(n^*) = \frac{q}{\sqrt{2}}\left(\frac{\gamma - 1}{\gamma}\right)\left(\frac{q}{2\gamma c\sigma\sqrt{Ark}}\right)^{1/(\gamma-1)}. \tag{16}$$

Recall that, in general, $\beta^* = b^*\sqrt{n/2}$, and n is decreasing in σ, c, r, and k. Since b^* is unaffected by exogenous parameters when $G(\cdot)$ is a power function, the combined effect is that β^* decreases in all of the parameters. The result seems counterintuitive, as one might expect the firm to raise incentives for the manager to monitor the workers more precisely if, for example, the measurement process worsens (σ rises). Instead, the firm shrinks the size of the team to respond to the increase in risk premium. Smaller teams require less monitoring and, hence, the firm can afford to reduce the incentives for the manager. The firm is once again using team size as the main instrument to respond to exogenous changes in the environment.

The findings in Propositions 2, 3, 5, and 6 provide a possible explanation for the lack of strong empirical support for the negative relation between risk and incentives suggested by traditional agency models (see Prendergast 2002), who suggests that firms may contract directly on employee effort in some settings, and contract on output in less certain settings. Under this construct, the strength of incentives given to employees need not vary monotonically with the level of uncertainty. Our model differs in that we assume that employee effort is unobservable and that contracts with employees can only be structured on observed joint output. In effect, while the Prendergast (2002) result points to the omission of delegation as a choice variable in traditional agency models, our finding is based on the ability of the principal to control team size in addition to employee incentives. To this extent, our model does in fact embrace a quite large class of organizations, since we consider any organizations where teams are key to the production process. In contrast, the existing empirical literature on the risk/incentive trade-off has looked only within a handful of narrow environments (e.g., executive compensation, sharecropping, sports tournaments).

Value of the Manager

Following our analyses of the mechanical monitoring and self-interested manager settings, we conclude this section by identifying when it is profitable for the firm to hire a manager. Suppose that firm size is exogenously given, and that $c_i = c$ for all $i \in N$. Assume the principal has the choice of either using a fixed monitor/information system or hiring a manager with known ability k. Then for any given set of parameters (q, r, c, σ^2), the firm compares the expected profit under the two alternatives. The value of hiring a manager is given by:

$$\Pi(n) - \hat{\Pi}(n) = \frac{n}{2c}\left[\left(q - 2c\sigma\sqrt{rkG(n)/n}\right)^2 - \frac{q^2}{1 + rcG(n)\sigma^2}\right]$$
$$= \frac{\sigma}{2}\sqrt{rG(n)}\left[q\sqrt{n}\left[\frac{q\sigma\sqrt{rnG(n)}}{1 + rcG(n)\sigma^2} - 4\sqrt{k}\right] + 4c\sigma\sqrt{rkG(n)}\right].$$

It follows immediately that a sufficient condition for the owner to find it in his interest to hire a manager with ability k is:

$$\frac{q\sigma\sqrt{rnG(n)}}{1 + rcG(n)\sigma^2} \geq 4\sqrt{k}. \tag{17}$$

As one might expect, the manager is valuable if his ability is sufficiently high: lower values of k imply fewer difficulties in motivating an effort- and risk-averse manager to work. In addition, the expression in Equation (17) shows that employing a manager is unambiguously worthwhile when the market value of output, q, is high. More interestingly, an analysis of the profit difference reveals that it is more worthwhile for the owner to hire a manager when the workers are of higher quality (i.e., c is lower). We explore more aspects of this complementarity between workers and managers in the next section.

The results with regard to the level of uncertainty in monitoring technology are less straightforward. In particular, it can be shown that the manager is not valuable when σ^2 is either too low (since a mechanical monitor can then do just as well) or too high (because of the high cost of motivating the manager and paying him a risk premium in that case). However, for all intermediate values of the uncertainty parameter, σ^2, the owner finds it beneficial to employ a manager. This suggests that if the precision of the performance measurement increases (i.e., σ decreases) over the life cycle of a firm, then the firm would operate without a monitoring manager initially, employ a monitoring manager when σ has decreased enough, and substitute in a mechanical monitor at a later, more mature stage of firm's life cycle when the performance measurement system is sufficiently precise.

When firm size is endogenous, then for a given set of parameters, the firm compares the expected profit under the mechanical monitor at the optimal team size and the expected profit under a hired manager at the (possibly different) optimal team size. While the economic intuition stated in the previous paragraph still holds, no general expression for the profit differential can be derived because of the implicit functions governing the optimal team sizes. We can, however, obtain additional insights by looking at the case where $G(n) = An^{\gamma}$. In this setting, the value of hiring a manager is:

$$\Pi(n^*) - \hat{\Pi}(\hat{n}) = \frac{q^2(\gamma - 1)}{2c\gamma}\cdot\left(\frac{n^*(\gamma - 1)}{\gamma} - \hat{n}\right).$$

As $\gamma > 1$, it follows that it is profitable to hire a manager if and only if the optimal team size with a manager exceeds the optimal team size without the manager. Managers assist in team performance measurement, and hence allow the formation of larger teams. The link between valuable managers and team sizes points to a synergy between the use of managers and the number of people assigned to them for supervision. What is interesting here is that the synergy (or complementarity) is not assumed of the production technology, but is due to the inherent moral hazard problem associated with team production.

To summarize, in this model, the firm has three design instruments to maximize its expected profit. They are: (1) the number of workers (n), (2) the choice of whether to hire a manager, and (3) the choice of incentive coefficients for the workers and manager (b and β). When the environment changes, as reflected by a variation in (q, σ, c, r, or k), all three design variables may change in response. The special property of the responses is a certain robustness in the worker-incentives; in particular, b^* responds less, and sometimes not at all, to changes other than in the market price of output, q, relative to the case when team size is fixed. For changes in all other environmental variables, the firm reacts more by changing team size and monitoring levels, and less by changing workers' incentives.

III. COMPLEMENTARITY BETWEEN MANAGERS AND WORKERS

In the previous section, we demonstrated the presence of endogenous synergies between team size and managerial supervision. In this section, we explore more such aspects of optimal incentives and monitoring in team settings. Of specific interest to us is the link between worker and manager quality. Recall that the parameters c and k in our model reflect the costs of effort for the workers and the manager; better workers and managers have lower c's and k's. This section explores complementarity, in the sense of optimal matching, between managers and workers. That is, should better managers be asked to supervise better workers or worse ones? We show that it is always optimal to do the former, i.e., manager and worker abilities are complements. We then explore how the selection of the optimal manager and worker combination in a labor market varies with uncertainty and other parameters of our model. Throughout the section, we will restrict our attention to the case when $G(\cdot)$ is a power function, $G(n) = An^{\gamma}$, with $A > 0$, $\gamma > 1$.

Positive Sorting

When $G(\cdot)$ is a power function, the closed-form solutions for the workers' and manager's incentive contracts are given in Equations (15) and (16). The solutions for the other variables of interest are provided below:

$$n^* = \left(\frac{q}{2\gamma c\sigma\sqrt{Ark}}\right)^{2/(\gamma-1)}; \tag{18}$$

$$m^* = \frac{q\sigma}{2}\left(\frac{\gamma-1}{\gamma}\right)\sqrt{\frac{Ar}{k}}\left(\frac{q}{2\gamma c\sigma\sqrt{Ark}}\right)^{(\gamma+1)/(\gamma-1)}; \tag{19}$$

$$\Pi^* = \frac{n^*(b^*)^2}{2c} = \frac{q^2}{2c}\left(\frac{\gamma-1}{\gamma}\right)^2\left(\frac{q}{2\gamma c\sigma\sqrt{Ark}}\right)^{2/(\gamma-1)}. \tag{20}$$

From Equation (20), note that the overall profit function is of the form

$$\Pi^* = K_0 k^{-1/(\gamma-1)} c^{-(\gamma+1)/(\gamma-1)}, \tag{21}$$

where $K_0 > 0$ is independent of k and c. Taking derivatives of Equation (21) shows that:

$$\frac{\partial \Pi^*}{\partial c} < 0 \quad \text{and} \quad \frac{\partial \Pi^*}{\partial k} < 0. \tag{22}$$

That is, better workers and managers raise profits. What drives this result? It is tempting to argue that it is easier for better workers to deliver a given level of effort, so the firm reduces their incentives, thus reducing the risk premium payments and, hence, raising profits. Similarly, it seems plausible that because better managers are more effective at reducing variance, this will also lead to lower risk premiums and higher profits. However, this logic neglects the main point of this paper: that the firm is simultaneously adjusting the team size, and this can reverse the usual comparative statics. To see this, note that in equilibrium m^* is increasing in n^*, which in turn is decreasing in k. A better manager allows the firm to increase team size, which causes the manager to work harder in equilibrium. Recall that, in general:

$$\mathrm{Var}(w_i) = b_i^2 \frac{G(n)\sigma^2}{m} \qquad \text{and} \qquad \mathrm{Var}(\omega) = \beta^2 \frac{G(n)\sigma^2}{m}, \tag{23}$$

so more managerial effort will pull wage variance downward. But the larger team size counteracts this effect. In fact, it turns out that for any power function the variance-increasing team-size effects of c and k dominate their direct variance-decreasing managerial effort effects. Substituting the optimal b^* and β^* from Equations (15)–(16) and the optimal n^* and m^* from Equations (18)–(19) into Equation (23) shows that, for $G(n) = An^\gamma$:

$$\mathrm{Var}(w_i) = K_1 c^{-1} \qquad \text{and} \qquad \mathrm{Var}(\omega) = K_2 c^{-(\gamma+1)/(\gamma-1)} k^{-1/(\gamma-1)}, \tag{24}$$

where K_1, $K_2 > 0$ are constants independent of c and k. We see that better workers and managers (lower c's and k's) in fact increase wage variance. In particular, observe that the risk premium:

$$RP = \frac{nr}{2}\,\mathrm{Var}(w) + \frac{r}{2}\,\mathrm{Var}(\omega),$$

increases with better workers and managers as well. We know that n decreases in c and k by (18), while $\mathrm{Var}(w)$ decreases in c and $\mathrm{Var}(\omega)$ decreases in c and k by Equation (24).

But if the risk payments are increasing, how do profits rise? Remember that larger teams not only inflate the risk premium, but also earn more production revenue for the firm. The firm maximizes total surplus, which (for identical workers) is:

$$TS = n \cdot \{qe - C(e)\} - [C(m) + RP].$$

Suppose the firm hires a better manager. This does not change the worker's effort choice, so the term in braces is unchanged. As shown above, the manager works harder and the risk premium rises; as a result, the term in brackets rises, pulling down profits. Finally, the firm hires a larger team, so the revenue from production lifts profits upward. The net effect is that the benefits from higher production exceed the costs of more managerial and more risk payments, so profits overall rise. A similar argument holds for better workers.

Taking cross partial derivatives of Equation (21), we see that:

$$\Pi^*_{ck} \equiv \frac{\partial^2 \Pi^*}{\partial c \partial k} = \frac{\gamma + 1}{(\gamma - 1)^2} K_0 k^{-\gamma/(\gamma-1)} c^{-2\gamma/(\gamma-1)} > 0, \tag{25}$$

which gives positive sorting: it is optimal to assign better managers to better workers.

Precisely, the condition above shows that $\partial\Pi^*/\partial c$ increases in k and $\partial\Pi^*/\partial k$ increases in c. Now $-\partial\Pi^*/\partial c$ is the marginal return from hiring a better worker. So, as the quality of the manager improves (i.e., k falls), $-\partial\Pi^*/\partial c$ increases, so the marginal benefit from hiring a better worker rises. Said differently, better workers raise profits, and in fact raise profits more under better managers.

To see this visually, refer to Figure 1. The two curves represent profit as a function of k for different costs of worker effort. Observe that both curves are decreasing: as the quality of the manager improves (k decreases), the firm earns more profit. Equation (25) says that $\Pi^*_k(\cdot)$ increases in c. So as c_2 rises to c_1, $\Pi^*_k(c_2, k)$ rises to $\Pi^*_k(c_1, k)$. Since $\Pi^*_k(\cdot) < 0$, this means $\Pi^*_k(c_1, k)$ is less negative than $\Pi^*_k(c_2, k)$, i.e., the profit function is flatter in k when $c = c_1$. In other words, as k_1 falls to k_2, profits will rise. But the increase in profit under the better worker (Δ_2) exceeds the increase in profit under the worse worker (Δ_1). So as the firm hires a better manager, the firm earns more profit by hiring a better worker as well.

This shows that good managers and good workers are complements. The intuition once again lies in the optimal team size. From Equation (18), it follows that the cross partial of n^* with respect to c and k is positive. With better workers and better managers, the firm can afford to increase team size; team size increases the most with good workers working under good managers. A good manager can more easily reduce the variance for a larger team, and even more so if the team consists of good workers. This complementarity between manager and worker quality at the team level translates to a complementarity at the profit level.

Optimal Manager and Agent Quality

The previous subsection shows that the firm earns more profits by hiring better workers in conjunction with better managers. This begs the question: Why not hire the best workers and monitor possible? The reason, of course, is that employing high-quality talent is expensive. We have, as does virtually all of the compensation/agency literature, thus far assumed a partial equilibrium setting where the labor market is captured only via the assumed (exogenous) reservation wage. In this section, we introduce the idea of a labor market by explicitly specifying the reservation wage as a function of labor talent.[11] This enables us to draw inferences about the impact of firm-specific factors, as well as worker and manager characteristics, on the optimal team size and labor mix employed by organizations.

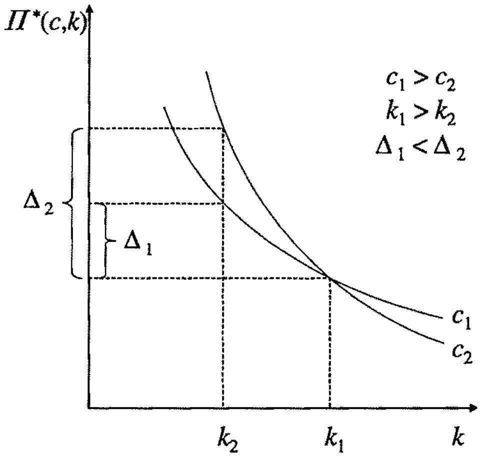
To model this in a parsimonious manner, we specify that the firm hires the manager and workers in distinct external labor markets. In particular, suppose that the quality $k > 0$ of the manager and the quality $c > 0$ of the worker are observable by both the firm and the market. A manager of quality k receives market wage $\bar{u}_M(k)$. Better managers (lower k) earn higher market wages, so $\bar{u}_M(k)$ is a positive and strictly decreasing function of k. Similarly, each identical worker receives a market wage of $\bar{u}_W(c)$, a positive and decreasing function of c.

It is clear that the firm must, and will, pay exactly these market wages to acquire its labor. As such, the firm sets the salaries (a, α) such that the workers' and manager's expected utilities precisely equal their market wages. Given binding (IR) constraints, the firm's profit function for exogenous n is then given by:

---

[11] Other papers that have modeled an outside labor market explicitly include Sappington (1983) and Ricart i Costa (1988). Huddart and Liang (2005) implicitly consider an outside labor market by characterizing partnerships that optimize per-partner utility, thereby making the reservation utility for any partner endogenous.

FIGURE 1
Positive Sorting across Managers and Workers



The curves represent profit as a function of manager quality (k) for different levels of worker quality (c). Higher levels of managerial or worker quality are indicated by lower values of k or c, respectively. The curves denote positive sorting as the marginal benefit (Δ) from having a better manager ($k_2 < k_1$) is greater when the worker is also of higher quality ($c_2 < c_1$).

$$\Pi(n, c, k|\sigma, r) = \frac{nb^2}{2c} - n \cdot \bar{u}_W(c) - \bar{u}_M(k),$$

where $b = b^* = q - 2c\sigma\sqrt{rkG(n)/n}$ is the optimal incentive payment for each worker (see Equation (13)).

Hiring a better quality manager (decreasing k) allows the firm to increase the size of the team, and hence generate more revenue from production. At the same time, better managers are more expensive, since $\bar{u}_M(k)$ increases as k decreases. The optimal k trades off the increased revenue from larger teams against the increased cost of a better quality manager, where this cost reflects the manager's market wage. A similar trade-off holds for worker quality. The firm chooses the optimal worker and manager quality in addition to optimal team size by solving the following program:

$$\max_{n,c,k} \Pi(n, c, k|\sigma, r) \text{ such that } n, c, k > 0, \text{ and } b \geq 0. \tag{26}$$

We have the following result.

Proposition 7: When $G(\cdot)$ is a power function, the optimal choices of b and c do not depend on σ or r.

The invariance property thus holds to an even greater extent in this setting. For exogenous changes in uncertainty or worker risk aversion, the firm optimally adjusts the team size and the manager quality, but keeps worker incentives and quality fixed. So invariance applies not just to incentives but to the quality of the worker as well. The inclusion of additional choice variables in the firm's problem allows the firm to use these instruments to adjust to changes in its environment. Essentially, the team size and manager quality are more flexible instruments in making the adjustments, compared to strength of incentives and worker quality.

If the firm does not alter b or c after changes in σ or r, how does it change n and k? As uncertainty or risk aversion increases, one might expect the firm to hire a better manager, since better managers by definition are more effective at reducing output variance. Instead, we have the following result:

Proposition 8: Suppose $G(\cdot)$ is a power function. As uncertainty (σ) or risk aversion (r) increases, it is optimal to decrease the team size, hire a lower quality manager, and keep the quality of the workers fixed.

Once again, it is important to consider that in equilibrium the firm adjusts the team size optimally. As shown earlier, the firm will shrink the team in response to higher uncertainty. Because managers are costly, it is a waste of resources to assign an expensive, high-quality manager to a small team when a lower quality manager will suffice. Even though the expensive manager by himself can more effectively reduce output variance than a cheaper manager, the firm uses team size as the primary instrument to deal with the increase in uncertainty, and then selects the quality of the manager to fit the team.

When $G(\cdot)$ is a power function, the parameter γ measures the quality of the performance technology. To see this, observe that when $G(n) = An^{\gamma}$, the output variance is:

$$\mathrm{Var}(y) = \frac{G(n)}{m}\sigma^2 = \frac{A\sigma^2}{m/n^{\gamma}}.$$

As γ increases, the manager's effectiveness of monitoring (m/n^{γ}) decreases, and hence the output variance increases. In other words, γ captures how quickly signal quality degrades as team size increases. The firm must of course bear the cost of this via higher risk payments. The next result characterizes how changes in the γ parameter affect the firm's choice of manager and worker quality.

Proposition 9: Suppose $G(\cdot)$ is a power function. As the performance measurement technology improves (γ decreases), it is optimal to enlarge the team size, hire better quality managers, and hire lower quality workers.

Note that any given manager is more effective at reducing variance if γ is low. Naturally, more effective managers can handle a larger number of workers, so the firm expands the team size. But surprisingly, a change in γ affects the optimal quality of the workers and the manager in opposite directions: the firm will hire lower quality workers but better quality managers. Again, focus on the effects coming from the optimal team size. As the team size increases, the firm hires a better manager to handle the larger team. In this sense, the quality of the performance measurement system and the quality of the manager are complementary. Better managers are more effective at utilizing a high-quality performance measurement system than managers of lower ability.

We also find that the firm hires lower quality workers. This arises from the combination of two forces: first, high-quality workers are now costly, and second, the firm is already earning more revenue from larger teams. Therefore, the firm expands the size of the team but lowers the quality of each team member. So team size and worker quality are substitutes in the firm's profit function, whereas team size and manager quality are complements. This arises directly from the nature of the monitoring production technology.

A noteworthy feature of the preceding comparative statics is that they require virtually no restrictions on the outside option functions. It is necessary only to impose that the functions are positive and strictly decreasing, but not to specify their functional form. The mere imposition of a cost of worker and manager quality suffices to generate rich comparative statics on the parameters of the model.

IV. CONCLUSION

In this paper, we study three instruments in the internal design of an organization involving team production: team size, monitoring activities, and the incentive contracts offered to workers and managers. When both types of agents (workers and managers) are self-interested and have the ability to take hidden actions, there exist complex interactions in the trade-offs among these instruments. We show, however, that such complex interactions produce rather simple and stark implications.

One such result is the robustness of worker contracts. We characterize the properties of attenuation and invariance whereby, compared to settings where team size is fixed, the equilibrium pay-for-performance scheme offered to workers varies less, if at all, with respect to most of the environmental variables of interest, including the extent of uncertainty, costs of effort, and risk preferences. By contrast, the incentive schemes offered to managers exhibit substantially higher levels of sensitivity to these parameters. Our work thus yields a clear prediction about the relative variation in the pay-for-performance sensitivities of employees in production versus supervisory roles in organizations.

We also characterize the optimal size of worker teams in various settings and provide testable predictions about the impact of environmental variables on the choice of optimal team size and structure, even in the presence of an external marketplace for talent. Of particular note, our work suggests the following implication: in settings with relatively lower levels of intrinsic uncertainty and/or better technology for measuring performance, team sizes will be higher and the quality of the workers in the teams will be either unchanged or lower, while the managers hired to oversee them will have greater levels of skill. The first of these implications is consistent with the finding by Rajan and Wulf (2006) of a significant flattening in the corporate hierarchy and a move to greater spans of control over the 1986–1998 time period, a shift they ascribe in part to improvements in information technology. More generally, Guadalupe and Wulf (2007) note that there is relatively little empirical evidence regarding the relationship between information technology and the structure of the internal hierarchy. The results in our paper provide several hypotheses that may be tested by future empirical research in this area.

Our theoretical exploration may be extended in at least three ways. First, we restrict attention to settings in which a single self-interested manager is responsible for monitoring a team of workers. It would be of great interest to extend the model framework to study the interaction between multiple managers who are asked to monitor a given worker team. Given the findings in Christensen and Sabac (2002), it is likely that the correlation in the managers' signals will prove pivotal in determining the value of having supervisory teams, and their impact on the optimal team size. This analysis could be expanded further, under alternative assumptions regarding other types of monitoring synergies, to explore ideas related to shirking in monitoring tasks or the assignment of managers to monitor different subgroups within a team. Another promising avenue for future study is the extension of our model to a multi-period setting. This would allow for analysis of the trade-offs inherent in team incentives and monitoring in a dynamic setting, and under various scenarios involving the firm's ability to commit to long-term incentive structures. Finally, all of our work has been carried out in a multi-agent LEN framework. We use this specification because it would be virtually impossible to replicate our analysis in a general agency model. However, it would be of interest to examine (perhaps via numerical simulations) whether our comparative static results, for example, hold in a broader range of settings.

APPENDIX

Proof of Proposition 1: Let v denote any of the symbols r, c, σ^2. Let $\bar{v}$ denote the product of the other two (i.e., if $v = r$, then $\bar{v} = c\sigma^2$). Implicitly differentiating (4) in v yields:

$$\frac{\partial \hat{n}}{\partial v} = -\frac{1}{\bar{v}\hat{n}G''(\hat{n})v^2} < 0. \tag{27}$$

■

Proof of Proposition 2: As before, let v denote any of the symbols r, c, σ^2 and let $\bar{v}$ denote the product of the other two. Differentiating (3) and applying (27), we obtain:

$$\delta^v_{\text{ex}}(\hat{n}) = \left.\frac{\partial \hat{b}}{\partial v}\right|_{n=\hat{n}} = -\frac{q\bar{v}G(\hat{n})}{(1 + rc\sigma^2 G(\hat{n}))^2} < 0; \tag{28}$$

$$\delta^v_{\text{end}}(\hat{n}) = \left.\frac{\partial \hat{b}}{\partial v}\right|_{n=\hat{n}} + \left.\frac{\partial \hat{b}}{\partial n}\right|_{n=\hat{n}} \cdot \frac{\partial \hat{n}}{\partial v} = -\frac{q\bar{v}}{(1 + rc\sigma^2 G(\hat{n}))^2}\left[G(\hat{n}) - \frac{G'(\hat{n})}{rc\sigma^2\hat{n}G''(\hat{n})}\right]. \tag{29}$$

From (4), we can substitute for $\frac{1}{rc\sigma^2}$ into the latter expression to obtain:

$$\begin{aligned} \delta^v_{\text{end}}(\hat{n}) &= -\frac{q\bar{v}}{(1 + rc\sigma^2 G(\hat{n}))^2}\left[G(\hat{n}) - \frac{G'(\hat{n})}{\hat{n}G''(\hat{n})} \cdot [\hat{n}G'(\hat{n}) - G(\hat{n})]\right] \\ &= -\frac{q\bar{v}G(\hat{n})}{(1 + rc\sigma^2 G(\hat{n}))^2}\left[\frac{G'(\hat{n})}{G''(\hat{n})} \cdot \left(\frac{G''(\hat{n})}{G'(\hat{n})} - \frac{G'(\hat{n})}{G(\hat{n})} + \frac{1}{\hat{n}}\right)\right] \leq 0, \end{aligned} \tag{30}$$

where the last weak inequality follows from the convexity of $G(\cdot)$ and Condition (1). From (28) and (30), it follows that $\delta^v_{\text{end}}(\hat{n}) > \delta^v_{\text{ex}}(\hat{n})$ if and only if:

$$\frac{G'(\hat{n})}{G''(\hat{n})} \cdot \left(\frac{G''(\hat{n})}{G'(\hat{n})} - \frac{G'(\hat{n})}{G(\hat{n})} + \frac{1}{\hat{n}}\right) < 1,$$

which reduces to: $\hat{n}G'(\hat{n}) - G(\hat{n}) > 0$, which holds for all $\hat{n} > 0$ since $G(0) = 0$ and $G(\cdot)$ is convex. ■

Proof of Proposition 3: For the "if" part of the result, suppose that $G(n) = An^\gamma$, $A > 0$, $\gamma > 1$. Then, for arbitrary n, the expression in square brackets in (30) simplifies to:

$$1 - \frac{G'(n)^2}{G(n)G''(n)} + \frac{G'(n)}{nG''(n)} = 1 - \frac{A^2\gamma^2 n^{2\gamma-2}}{(An^\gamma)\cdot(A\gamma(\gamma-1)n^{\gamma-2})} + \frac{A\gamma n^{\gamma-1}}{nA\gamma(\gamma-1)n^{\gamma-2}}$$
$$= 1 - \frac{\gamma}{\gamma-1} + \frac{1}{\gamma-1} = 0. \tag{31}$$

Therefore, $\delta^v_{\text{end}}(n) = 0$ for all values of n, i.e., pay-for-performance sensitivity is unaffected by changes to the parameter values (σ^2, r, c) once team size is adjusted optimally.

For necessity, note from Proposition 1 that changes to σ^2, r, or c lead to variations in the optimal n via Equation (4). Therefore, if b is invariant to the parameters of the problem, then it must be the case that $\delta^v_{\text{end}}(n) = 0$ for an open set of values of n. From (30), this implies that for all n in some interval:

$$\frac{G''(n)}{G'(n)} - \frac{G'(n)}{G(n)} + \frac{1}{n} = 0$$
$$\Leftrightarrow \frac{d}{dn}\left(\ln\left[\frac{nG'(n)}{G(n)}\right]\right) = 0$$
$$\Rightarrow \frac{nG'(n)}{G(n)} = \theta,$$
$$\Rightarrow G'(n) - \frac{\theta}{n}\cdot G(n) = 0 \tag{32}$$

for some constant $\theta > 0$. Solving the first-order linear differential equation in (32) reveals that $G(\cdot)$ must belong to the class of power functions, thus completing the proof. ■

Proof of Lemma 1: From Equations (13) and (14), the owner's profit function is:

$$\Pi(n) = \frac{n}{2c}(q - 2c\sigma\sqrt{rkG(n)/n})^2, \tag{33}$$

and this is defined over the relevant range where $b^*(n) = q - 2c\sigma\sqrt{rkG(n)/n} \geq 0$. Note that $h(n) = \sqrt{\frac{G(n)}{n}}$ is an increasing function of n. This implies $n \in [0, \bar{n}]$, where $\bar{n}$ satisfies: $q - 2c\sigma\sqrt{rk}h(\bar{n}) = 0$. Evaluating $\Pi(n)$ at the two bounds, it is clear that $\Pi(\bar{n})$. Moreover:

$$\Pi(0) = \frac{1}{2c}\cdot\lim_{n\to 0}(nq^2 + 4c^2\sigma^2 rkG(n) - 4qc\sigma\sqrt{rknG(n)}) = 0.$$

By Rolle's theorem, we are therefore assured that an interior n^* exists with $\Pi'(n^*) = 0$. Now, the derivative of $\Pi(n)$ with respect to team size is given by:

$$\Pi'(n) = \frac{b^*(n)}{2c}(q - 2c\sigma\sqrt{rk}h(n) - 4nc\sigma\sqrt{rk}h'(n)).$$

Since $b^*(n) > 0$ in the interior, any n^* that satisfies $\Pi'(n^*) = 0$ is characterized by:

$$h(n^*) + 2n^*h'(n^*) = \frac{q}{2c\sigma\sqrt{rk}}. \tag{34}$$

The expression on the left is an increasing function of n if and only if $3h'(n) + 2nh''(n) > 0$. As $h(n) \equiv \sqrt{\frac{G(n)}{n}}$, we have:

$$h'(n) = \frac{\sqrt{n}G'(n)}{2\sqrt{G(n)}}\left(\frac{1}{n} - \frac{G(n)}{G'(n)n^2}\right) \tag{35}$$

and:

$$h''(n) = \frac{\sqrt{n}G'(n)}{4n\sqrt{G(n)}}\left(\frac{2G''(n)}{G'(n)} + \frac{3G(n)}{n^2G'(n)} - \frac{G'(n)}{G(n)} - \frac{2}{n}\right). \tag{36}$$

Together, (35) and (36) imply that:

$$\begin{aligned} 3h'(n) + 2nh''(n) &= \frac{\sqrt{n}G'(n)}{2\sqrt{G(n)}} \cdot \left(2\frac{G''(n)}{G'(n)} - \frac{G'(n)}{G(n)} + \frac{1}{n}\right) \\ &> \frac{\sqrt{n}G'(n)}{2\sqrt{G(n)}} \cdot \left(\frac{G''(n)}{G'(n)} - \frac{G'(n)}{G(n)} + \frac{1}{n}\right) \geq 0, \end{aligned} \tag{37}$$

where the strict and weak inequalities follow from $G''(\cdot) > 0$ and Condition (1), respectively. Therefore, a unique n^* exists satisfying (34). To verify the second-order conditions at this point, we differentiate $\Pi'(n)$ with respect to n to obtain:

$$\begin{aligned} \Pi''(n) &= \frac{b^{*\prime}(n)}{2c}(q - 2c\sigma\sqrt{rk}h(n) - 4nc\sigma\sqrt{rk}h'(n)) \\ &\quad + \frac{b^*(n)}{2c}(-2c\sigma\sqrt{rk}h'(n) - 4c\sigma\sqrt{rk}h'(n) - 4nc\sigma\sqrt{rk}h''(n)). \end{aligned}$$

Evaluating this expression at n^*, we have:

$$\Pi''(n^*) = 0 + \frac{b^*(n^*)}{2c}(-2c\sigma\sqrt{rk})(3h'(n^*) + 2n^*h''(n^*)) < 0,$$

where the last inequality follows from (37) and $b^*(n^*) > 0$. We can therefore conclude that n^* is the globally optimal team size. ■

Proof of Proposition 4: Let v denote any of c, σ, $\sqrt{r}$, and $\sqrt{k}$, and let $\bar{v}$ denote the product of the other three. Implicitly differentiating the condition for n^* in (34) shows that:

$$\frac{\partial n^*}{\partial v} = (3h'(n^*) + 2n^*h''(n^*))^{-1}\left(-\frac{q}{2\bar{v}v^2}\right) < 0,$$

as the first term is strictly positive from (37). ■

Proof of Proposition 5: As before, let v denote any of c, σ, $\sqrt{r}$, and $\sqrt{k}$, and let $\bar{v}$ denote the product of the other three. First, differentiating (13), we immediately obtain:

$$\varphi^{v}_{ex}(n^*) = -2\bar{v}h(n^*) < 0. \tag{38}$$

Next, implicitly differentiating the first order-Condition (34) in Lemma 1, we have:

$$\frac{\partial n^*}{\partial v} = -\frac{q\bar{v}}{2(3h'(n^*) + 2nh''(n^*))(v\bar{v})^2} < 0 \text{ (from (37))}.$$

We can then compute $\varphi^{v}_{end}(n^*)$ as:

$$\begin{aligned}\varphi^{v}_{end}(n^*) &= \left.\frac{\partial b^*}{\partial v}\right|_{n=n^*} + \left.\frac{\partial b^*}{\partial n}\right|_{n=n^*} \cdot \left.\frac{\partial n^*}{\partial v}\right|_{n=n^*} = -2\bar{v}h(n^*) \\ &\quad + \frac{2q\bar{v}h'(n^*)}{(3h'(n^*) + 2n^*h''(n^*))(2v\bar{v})}.\end{aligned}$$

Substituting for $q/(2v\bar{v})$ using Condition (34), we obtain:

$$\varphi^{v}_{end}(n^*) = -2\bar{v}\left[h(n^*) - \frac{h'(n^*)(h(n^*) + 2n^*h'(n^*))}{3h'(n^*) + 2n^*h''(n^*)}\right] \tag{39}$$

$$= -4\bar{v}\left[\frac{h'(n^*)h(n^*) + n^*h(n^*)h''(n^*) - n^*h'(n^*)^2}{3h'(n^*) + 2n^*h''(n^*)}\right]. \tag{40}$$

The denominator of (40) is strictly positive (from (37)). Using Equations (35) and (36), the numerator can be simplified to:

$$\frac{G'(n)}{2} \cdot \left(\frac{G''(n)}{G'(n)} - \frac{G'(n)}{G(n)} + \frac{1}{n}\right) \geq 0, \tag{41}$$

where the inequality follows from Condition (1).

We have therefore established that $\varphi^{v}_{end}(n^*) \leq 0$. Moreover, using (37) and the fact that $h'(\cdot) > 0$, a direct comparison of (38) and (39) reveals that $\varphi^{v}_{ex}(n^*) < \varphi^{v}_{end}(n^*) \leq 0$.

Turning to the manager's incentives, we know that with identical workers:

$$\beta^* = b^* \sqrt{\frac{n}{2}}.$$

Therefore:

$$\psi^{v}_{ex}(n^*) = \sqrt{\frac{n^*}{2}} \cdot \varphi^{v}_{ex}(n^*) = -2\bar{v}h(n^*)\sqrt{\frac{n^*}{2}} < 0. \tag{42}$$

A comparison of (38) and (42) immediately reveals that $\psi^{v}_{ex}(n^*) \leq \varphi^{v}_{ex}(n^*)$, with strict inequality for all $n^* > 2$.

To complete the proof, we next demonstrate that $\psi^{\nu}_{\text{ex}}(n^*) = \psi^{\nu}_{\text{end}}(n^*)$. Since:

$$\psi^{\nu}_{\text{ex}}(n^*) = \psi^{\nu}_{\text{end}}(n^*) + \left.\frac{\partial\beta^*}{\partial n}\right|_{n=n^*} \cdot \left.\frac{\partial n^*}{\partial \nu}\right|_{n=n^*},$$

it is sufficient to show that $\left.\frac{\partial\beta^*}{\partial n}\right|_{n=n^*} = 0$. Using (13), we can expand β^* as follows:

$$\beta^* = b^* \sqrt{\frac{n}{2}} = \sqrt{\frac{n}{2}} \cdot (q - 2c\sigma\sqrt{rkG(n)/n}) = \sqrt{\frac{n}{2}} \cdot (q - 2\nu\bar{\nu}\sqrt{G(n)/n}).$$

Differentiating this expression, we obtain:

$$\begin{aligned}\left.\frac{\partial\beta^*}{\partial n}\right|_{n=n^*} &= \frac{q}{\sqrt{2}}\frac{1}{2\sqrt{n^*}} - \frac{\nu\bar{\nu}}{\sqrt{2}}\frac{G'(n^*)}{\sqrt{G(n^*)}} = \frac{\nu\bar{\nu}}{\sqrt{2n^*}} \cdot \left(\frac{q}{2\nu\bar{\nu}} - \frac{G'(n^*)\sqrt{n^*}}{\sqrt{G(n^*)}}\right)\\ &= \frac{\nu\bar{\nu}}{\sqrt{2n^*}} \cdot \left(h(n^*) + 2nh'(n^*) - \frac{G'(n^*)\sqrt{n^*}}{\sqrt{G(n^*)}}\right) \text{ (from (34))}\\ &= \frac{\nu\bar{\nu}}{\sqrt{2n^*}} \cdot \left(\sqrt{\frac{G(n^*)}{n^*}} + \frac{\sqrt{n^*}G'(n^*)}{\sqrt{G(n^*)}} - \frac{\sqrt{G(n^*)}}{\sqrt{n^*}} - \frac{G'(n^*)\sqrt{n^*}}{\sqrt{G(n^*)}}\right) = 0.\end{aligned}$$

■

Proof of Proposition 6: Suppose $G(n) = An^{\gamma}$, $A > 0$, $\gamma > 1$. Then, for arbitrary n^*, the numerator of $\varphi^{\nu}_{\text{end}}(n^*)$, given in reduced form in (41), simplifies to:

$$\begin{aligned}\frac{G'(n^*)}{2} \cdot \left(\frac{G''(n^*)}{G'(n^*)} - \frac{G'(n^*)}{G(n^*)} + \frac{1}{n^*}\right) &= \frac{G'(n^*)}{2}\\ &\cdot ((\gamma - 1)(n^*)^{-1} - \gamma(n^*)^{-1} + (n^*)^{-1}) = 0.\end{aligned}$$

Thus, $\varphi^{\nu}_{\text{end}}(n^*) = 0$ everywhere, i.e., pay-for-performance is totally insensitive to changes in any of the four parameters of interest, (σ^2, r, c, k), provided team size is adjusted optimally.

For the "only if" argument, we have shown in Proposition 5 that changes to σ^2, r, c, or k lead to strict variations in the optimal n via Equation (34). Therefore, if b is invariant to any of these parameters, it implies that $\varphi^{\nu}_{\text{end}}(n) = 0$ for an open set of values of n. From Equations (39)–(41), this implies that for all n in some open interval:

$$\frac{G''(n)}{G'(n)} - \frac{G'(n)}{G(n)} + \frac{1}{n} = 0.$$

As in the proof of Proposition 3, this implies that $G(\cdot)$ belongs to the class of power functions. ■

Proof of Proposition 7: When $G(n)$ is a power function, we can write $h(n) = \sqrt{G(n)/n} = Bn^{\tau}$, where B, τ are arbitrary, positive real numbers. The objective function (see (26)) is then:

$$\Pi(n, c\ k; \sigma, r) = \frac{n(q - 2c\sigma\sqrt{rk}Bn^{\tau})^2}{2c} - n \cdot \bar{u}_W(c) - \bar{u}_M(k). \tag{43}$$

If we use $b = q - 2c\sigma\sqrt{rk}Bn^{\tau}$ instead of n (change of variables), then:

$$n = \left(\frac{1}{B\sigma\sqrt{rk}}\right)^{1/\tau}\left(\frac{q - b}{2c}\right)^{1/\tau}.$$

Substituting into (43) and getting rid of n, we obtain the objective function:

$$\hat{\Pi}(b, c, k; \sigma, r) = \left(\frac{1}{B\sigma\sqrt{rk}}\right)^{1/\tau}\left(\frac{q - b}{2c}\right)^{1/\tau}\left(\frac{b^2}{2c} - \bar{u}_W(c)\right) - \bar{u}_M(k).$$

The optimal b and c are then determined from the following optimization:

$$\max_{b,c} \hat{\Pi}(b, c, k; \sigma, r) \text{ such that } 0 \leq b < q \text{ and } c > 0.$$

Obviously, this problem is equivalent to:

$$\max_{b,c} \left(\frac{q - b}{2c}\right)^{1/\tau}\left(\frac{b^2}{2c} - \bar{u}_W(c)\right)$$

such that $0 \leq b < q$ and $c > 0$, which does not depend on σ or r. ■

Proof of Proposition 8: Expanding the objective function in (43), we obtain:

$$\begin{aligned}\Pi(n, c, k; \sigma, r) &= \frac{q^2}{2c}n - 2q\sigma\sqrt{rk}Bn^{\tau+1} + 2c\sigma^2 rkB^2n^{2\tau+1} \\ &\quad - n \cdot \bar{u}_W(c) - \bar{u}_M(k).\end{aligned} \tag{44}$$

The maximization problem now becomes:

$$\max_{n,c,k} \Pi(n, c, k|\sigma, r) \text{ such that } n, c, k > 0 \text{ and } q - 2c\sigma\sqrt{rk}h(n) \geq 0.$$

First, observe that the comparative statics with respect to σ and r share the same sign. The optimal choice functions are $Z^*(\sigma, r)$ where $Z = n$, c or k. Note that parameters σ and r enter the problem only through the expression $\sigma\sqrt{r}$. This implies that $Z^*(\sigma, r) = Z^*(\sigma\sqrt{r}, 1)$. Therefore, $\operatorname{sign} \frac{\partial Z^*(\sigma, r)}{\partial \sigma} = \operatorname{sign} \frac{\partial Z^*(\sigma, r)}{\partial r} = \operatorname{sign} \frac{\partial Z^*(\sigma\sqrt{r}, 1)}{\partial \sigma}$.

Assumption 1: *The optimum is reached in the interior.*
First Order Conditions:

$$\Pi_n = \frac{q^2}{2c} - 2q(1 + \tau)\sigma\sqrt{rk}Bn^{\tau} + 2(1 + 2\tau)c\sigma^2 rkB^2 n^{2\tau} - \bar{u}_W(c) = 0; \tag{45a}$$

$$\Pi_c = n\left(-\bar{u}'_W(c) - \frac{q^2}{2c^2} + 2\sigma^2 rkB^2 n^{2\tau}\right) = 0; \tag{45b}$$

$$\Pi_k = -\bar{u}'_M(k) - q\sigma\sqrt{\frac{r}{k}}Bn^{\tau+1} + 2c\sigma^2 rB^2 n^{2\tau+1} = 0. \tag{45c}$$

The optimal values $n^*(\sigma, r)$, $c^*(\sigma, r)$, and $k^*(\sigma, r)$ solve system (45). The implicit function proposition ensures the existence of differentiable implicit functions $n^*(\sigma, r)$, $c^*(\sigma, r)$, and $k^*(\sigma, r)$, provided the second derivative matrix $D^2\Pi$ is nondegenerate.

Assumption 2: *Matrix $D^2\Pi$ is negative definite at the optimal solution.*

A matrix is negative definite if and only if the leading principal minors switch signs:

$$\Pi_{nn} < 0 \tag{46}$$

$$\Pi_{nn}\Pi_{cc} - \Pi_{nc}^2 > 0 \tag{47}$$

$$\det D^2\Pi < 0 \tag{48}$$

Equation (48) shows that we can apply the implicit function theorem to system (45). Differentiating (45) w.r.t. σ, we get:

$$\begin{pmatrix} \partial n^*/\partial\sigma \\ \partial c^*/\partial\sigma \\ \partial k^*/\partial\sigma \end{pmatrix} = -(D^2\Pi)^{-1}\begin{pmatrix} \Pi_{n\sigma} \\ \Pi_{c\sigma} \\ \Pi_{k\sigma} \end{pmatrix}. \tag{49}$$

Second derivatives. Differentiating (45a), we get:

$$\Pi_{nn} = 4\tau(1 + 2\tau)c\sigma^2 rkB^2 n^{2\tau-1} - 2q\tau(\tau + 1)\sigma\sqrt{rk}Bn^{\tau-1}. \tag{50}$$

Differentiating (45b) with respect to n and using the FOC for c, we get:

$$\Pi_{nc} = -\bar{u}'_W(c) - \frac{q^2}{2c^2} + 2\sigma^2 rkB^2 n^{2\tau} + 4\tau\sigma^2 rkB^2 n^{2\tau} = 4\tau\sigma^2 rkB^2 n^{2\tau}, \tag{51}$$

where $-\bar{u}'_W(c) - \frac{q^2}{2c^2} + 2\sigma^2 rkB^2 n^{2\tau} = 0$ by (45b). Differentiating (45c) with respect to n, we get:

$$\Pi_{nk} = -q(\tau + 1)\sigma\sqrt{\frac{r}{k}}Bn^{\tau} + 2(2\tau + 1)\sigma^2 crB^2 n^{2\tau}.$$

In a similar manner:

$$\begin{aligned}
\Pi_{cc} &= n\left(-\bar{u}''_W(c) + \frac{q^2}{c^3}\right),\\
\Pi_{ck} &= 2\sigma^2 rB^2 n^{2\tau-1},\\
\Pi_{kk} &= -\bar{u}''_M(k) + \frac{1}{2}\, q\sigma r^{1/2} k^{-3/2} Bn^{\tau+1},\\
\Pi_{n\sigma} &= -2q(1+\tau)\sqrt{rk}Bn^{\tau} + 4(1+2\tau)c\sigma rkB^2 n^{2\tau},\\
\Pi_{c\sigma} &= 4\sigma rkB^2 n^{2\tau+1},\\
\Pi_{k\sigma} &= -q\sqrt{\frac{r}{k}}\, Bn^{\tau+1} + 4c\sigma rB^2 n^{2\tau+1},\\
\Pi_{nr} &= -q(1+\tau)\sigma\sqrt{\frac{k}{r}}\, Bn^{\tau} + 2(1+2\tau)c\sigma^2 kB^2 n^{2\tau},\\
\Pi_{cr} &= 2\sigma^2 kB^2 n^{2\tau+1},\\
\Pi_{kr} &= -q\sigma\frac{1}{2\sqrt{rk}}\, Bn^{\tau+1} + 2c\sigma^2 B^2 n^{2\tau+1}.
\end{aligned} \tag{52}$$

Finally, by direct calculations, using the formulas above, we can see that:

$$\Pi_{n\sigma} = \frac{n}{\tau\sigma}\Pi_{nn}, \quad \Pi_{nk} = \frac{n}{2\tau k}\Pi_{nn}. \tag{53}$$

Similarly:

$$\Pi_{c\sigma} = \frac{n}{\tau\sigma}\Pi_{nc}, \quad \Pi_{ck} = \frac{n}{2\tau k}\Pi_{nc}. \tag{54}$$

Behavior of $c^*(\sigma, r)$. Applying Cramer's rule to the linear system (49), we get:

$$\begin{aligned}
-\det D^2\Pi\,\frac{\partial c^*}{\partial\sigma} &= -\Pi_{n\sigma}(\Pi_{nc}\Pi_{kk} - \Pi_{ck}\Pi_{nk})\\
&+ \Pi_{c\sigma}(\Pi_{nn}\Pi_{kk} - (\Pi_{nk})^2) - \Pi_{k\sigma}(\Pi_{nn}\Pi_{ck} - \Pi_{nk}\Pi_{nc}).
\end{aligned}$$

After eliminating Π_{nk}, $\Pi_{n\sigma}$, $\Pi_{c\sigma}$ and Π_{ck} by substituting (53) and (54), we obtain:

$$\begin{aligned}
-\det D^2\Pi\,\frac{\partial c^*}{\partial\sigma} = {}& -\frac{n}{\tau\sigma}\Pi_{nn}\left(\Pi_{nc}\Pi_{kk} - \frac{n^2}{4\tau^2 k^2}\Pi_{nc}\Pi_{nn}\right)\\
&+ \frac{n}{\tau\sigma}\Pi_{nc}\left(\Pi_{nn}\Pi_{kk} - \frac{n^2}{4\tau^2 k^2}(\Pi_{nn})^2\right)\\
&-\Pi_{k\sigma}\left(\frac{n}{2\tau k}\Pi_{nn}\Pi_{nc} - \frac{n}{2\tau k}\Pi_{nn}\Pi_{nc}\right) = 0.
\end{aligned}$$

Now $\det D^2\Pi \neq 0$, from (48), so $\frac{\partial c^*}{\partial\sigma} = 0$.

Behavior of $k^*(\sigma, r)$. Again using Cramer's rule for the linear system (49):

$$\begin{aligned}-\det D^2\Pi \frac{\partial k^*}{\partial \sigma} &= \Pi_{n\sigma}(\Pi_{nc}\Pi_{ck} - \Pi_{cc}\Pi_{nk})\\ &-\Pi_{c\sigma}(\Pi_{nn}\Pi_{ck} - \Pi_{nc}\Pi_{nk}) + \Pi_{k\sigma}(\Pi_{nn}\Pi_{cc} - (\Pi_{nc})^2).\end{aligned}$$

We eliminate Π_{nk}, $\Pi_{n\sigma}$, $\Pi_{c\sigma}$, and Π_{ck} by substituting for (53) and (54):

$$\begin{aligned}-\det D^2\Pi \frac{\partial k^*}{\partial \sigma} &= \frac{n}{\tau\sigma}\Pi_{nn}\left(\frac{n}{2\tau k}(\Pi_{nc})^2 - \frac{n}{2\tau k}\Pi_{cc}\Pi_{nn}\right)\\ &\quad - \frac{n}{\tau\sigma}\Pi_{nc}\left(\frac{n}{2\tau k}\Pi_{nn}\Pi_{nc} - \frac{n}{2\tau k}\Pi_{nc}\Pi_{nn}\right) + \Pi_{k\sigma}(\Pi_{nn}\Pi_{cc} - (\Pi_{nc})^2)\\ &= (\Pi_{nn}\Pi_{cc} - (\Pi_{nc})^2)\left(\Pi_{k\sigma} - \frac{n^2}{2\tau^2 k\sigma}\Pi_{nn}\right). \end{aligned} \tag{55}$$

The last factor in the expression above is, after simplifying:

$$\Pi_{k\sigma} - \frac{n^2}{2\tau^2 k\sigma}\Pi_{nn} = \frac{Bn^{\tau+1}}{\tau}\sqrt{\frac{r}{k}}\,b. \tag{56}$$

Substituting back into (55), we get:

$$\frac{\partial k^*}{\partial \sigma} = -\frac{1}{\det D^2\Pi}(\Pi_{nn}\Pi_{cc} - (\Pi_{nc})^2)\frac{Bn^{\tau+1}}{\tau}\sqrt{\frac{r}{k}}\,b > 0,$$

because $\det D^2\Pi < 0$ and $\Pi_{nn}\Pi_{cc} - (\Pi_{nc})^2 > 0$, by (47) and (48).

Behavior of $n^*(\sigma, r)$. We use Cramer's rule for the linear system (49) to obtain:

$$\begin{aligned}-\det D^2\Pi \frac{\partial n^*}{\partial \sigma} &= \Pi_{n\sigma}(\Pi_{cc}\Pi_{kk} - (\Pi_{ck})^2)\\ &-\Pi_{c\sigma}(\Pi_{nc}\Pi_{kk} - \Pi_{nk}\Pi_{ck}) + \Pi_{k\sigma}(\Pi_{nc}\Pi_{ck} - \Pi_{cc}\Pi_{nk}).\end{aligned}$$

We eliminate Π_{nk}, $\Pi_{n\sigma}$, $\Pi_{c\sigma}$, and Π_{ck} by substituting (53) and (54). This yields:

$$-\det D^2\Pi \frac{\partial n^*}{\partial \sigma} = \left(\frac{n}{\tau\sigma}\Pi_{kk} - \frac{n}{2\tau k}\Pi_{k\sigma}\right)(\Pi_{nn}\Pi_{cc} - (\Pi_{nc})^2). \tag{57}$$

We first show that the first factor in the last expression above is negative. Since the ordering of the rows and columns of the Hessian $D^2\Pi$ is arbitrary, we can swap the second and third columns and rows of the matrix. As before, the matrix is negative definite if and only if the leading principal minors alternate in sign. This implies that $\Pi_{nn}\Pi_{kk} - (\Pi_{nk})^2 > 0$. Using (53):

$$0 < \Pi_{nn}\Pi_{kk} - (\Pi_{nk})^2 = \Pi_{nn}\left(\Pi_{kk} - \frac{n^2}{4\tau^2k^2}\Pi_{nn}\right). \tag{58}$$

As $\Pi_{nn} < 0$ from (46), inequality (58) implies:

$$\Pi_{kk} < \frac{n^2}{4\tau^2k^2}\Pi_{nn}. \tag{59}$$

Using this inequality to eliminate Π_{kk} from (57), and using (56) to simplify, we get

$$\begin{aligned}\frac{n}{\tau\sigma}\Pi_{kk} - \frac{n}{2\tau k}\Pi_{k\sigma} &< \frac{n}{\tau\sigma} \times \frac{n^2}{4\tau^2k^2}\Pi_{nn} - \frac{n}{2\tau k}\Pi_{k\sigma} \\ &= -\frac{n}{2\tau k}\frac{Bn^{\tau+1}}{\tau}\sqrt{\frac{r}{k}}\,b < 0.\end{aligned} \tag{60}$$

Thus, from (57) we get:

$$\frac{\partial n^*}{\partial \sigma} = -\frac{1}{\det D^2\Pi}\left(\frac{n}{\tau\sigma}\Pi_{kk} - \frac{n}{2\tau k}\Pi_{k\sigma}\right)(\Pi_{nn}\Pi_{cc} - (\Pi_{nc})^2) < 0,$$

where $\det D^2\Pi < 0$ by (48), $\frac{n}{\tau\sigma}\Pi_{kk} - \frac{n}{2\tau k}\Pi_{k\sigma} < 0$ by (60), and $\Pi_{nn}\Pi_{cc} - (\Pi_{nc})^2 > 0$ by (47). ■

Proof of Proposition 9: When $G(n) = An^\gamma$, $h(n) = \sqrt{G(n)/n} = Bn^\tau$, where $\tau = \frac{\gamma - 1}{2} > 0$. As such, comparative statics with respect to γ mirror those with respect to τ. We compute the latter, and start by differentiating (45) with respect to τ:

$$\begin{aligned}\Pi_{n\tau} = {} & (4(1 + 2\tau)c\sigma^2rkB^2n^{2\tau} - 2q(1 + \tau)\sigma\sqrt{rk}Bn^\tau)\left(\frac{1}{1 + 2\tau} + \ln n\right) \\ & - \frac{2\tau}{1 + 2\tau}q\sigma\sqrt{rk}Bn^\tau.\end{aligned}$$

Since $4(1 + 2\tau)c\sigma^2rkB^2n^{2\tau} - 2q(1 + \tau)\sigma\sqrt{rk}Bn^\tau = \frac{n}{\tau}\Pi_{nn}$ by (50):

$$\Pi_{n\tau} = \Pi_{nn}\frac{n}{\tau}\left(\frac{1}{1 + 2\tau} + \ln n\right) - \frac{2\tau}{1 + 2\tau}q\sigma\sqrt{rk}Bn^\tau. \tag{61}$$

The two remaining equations are straightforward:

$$\Pi_{c\tau} = 4\sigma^2 rkB^2 n^{2\tau+1} \ln n = \Pi_{nc}\frac{n}{\tau}\ln n \text{ (see Equation (51))} \tag{62}$$

$$\Pi_{k\tau} = \left(-q\sigma\sqrt{\frac{r}{k}}Bn^{\tau+1} + 4c\sigma^2 rB^2 n^{2\tau+1}\right)\ln n = \Pi_{k\sigma}\sigma \ln n \text{ (see Equation (52))} \tag{63}$$

Behavior of $c^*(\tau)$. By applying Cramer's rule to a system analogous to (49) (with τ in place of σ), we get:

$$\begin{aligned}-\det D^2\Pi\frac{\partial c^*}{\partial \tau} &= -\Pi_{n\tau}(\Pi_{nc}\Pi_{kk} - \Pi_{nk}\Pi_{ck}) \\ &\quad + \Pi_{c\tau}(\Pi_{nn}\Pi_{kk} - (\Pi_{nk})^2) - \Pi_{k\tau}(\Pi_{nn}\Pi_{ck} - \Pi_{nc}\Pi_{nk}).\end{aligned}$$

We use (53) and (54) to eliminate Π_{ck} and Π_{nk} and obtain:

$$-\det D^2\Pi\frac{\partial c^*}{\partial \tau} = (\Pi_{c\tau}\Pi_{nn} - \Pi_{nc}\Pi_{n\tau})\left(\Pi_{kk} - \frac{n^2}{4\tau^2k^2}\Pi_{nn}\right). \tag{64}$$

We transform the first factor in Equation (64), using (61) and (62) to eliminate $\Pi_{n\tau}$ and $\Pi_{c\tau}$:

$$\Pi_{c\tau}\Pi_{nn} - \Pi_{nc}\Pi_{n\tau} = \Pi_{nc}\left(-\Pi_{nn}\frac{n}{\tau}\frac{1}{1+2\tau} + \frac{2\tau}{1+2\tau}q\sigma\sqrt{rk}Bn^{\tau}\right) > 0, \tag{65}$$

since $\Pi_{nc} > 0$ by (51), and $-\Pi_{nn} > 0$ by (46). Finally, we get:

$$\frac{\partial c^*}{\partial \tau} = -\frac{1}{\det D^2\Pi}(\Pi_{c\tau}\Pi_{nn} - \Pi_{nc}\Pi_{n\tau})\left(\Pi_{kk} - \frac{n^2}{4\tau^2k^2}\Pi_{nn}\right) < 0,$$

since $\det D^2\Pi < 0$ by (48), $\Pi_{c\tau}\Pi_{nn} - \Pi_{nc}\Pi_{n\tau} > 0$ by (65), and $\Pi_{kk} - \frac{n^2}{4\tau^2k^2}\Pi_{nn} < 0$ by (59).

Behavior of $n^*(\tau)$. By applying Cramer's rule to a system analogous to (49) (with τ in place of σ), we get:

$$\begin{aligned}-\det D^2\Pi\frac{\partial n^*}{\partial \tau} &= \Pi_{n\tau}(\Pi_{cc}\Pi_{kk} - (\Pi_{ck})^2) \\ &\quad -\Pi_{c\tau}(\Pi_{nc}\Pi_{kk} - \Pi_{ck}\Pi_{nk}) + \Pi_{k\tau}(\Pi_{nc}\Pi_{ck} - \Pi_{cc}\Pi_{nk}).\end{aligned}$$

We use (53) and (54) to eliminate Π_{ck} and Π_{nk} from this equation to obtain:

$$\begin{aligned}-\det D^2\Pi\frac{\partial n^*}{\partial \tau} &= (\Pi_{nn}\Pi_{cc} - (\Pi_{nc})^2)\frac{n}{2\tau k}\left(\frac{n}{2\tau k}\Pi_{n\tau} - \Pi_{k\tau}\right) \\ &\quad + (\Pi_{n\tau}\Pi_{cc} - \Pi_{c\tau}\Pi_{nc})\left(\Pi_{kk} - \frac{n^2}{4\tau^2k^2}\Pi_{nn}\right),\end{aligned} \tag{66}$$

where $\Pi_{nn}\Pi_{cc} - (\Pi_{nc})^2 > 0$ by (47), and $\Pi_{kk} - \frac{n^2}{4\tau^2k^2}\Pi_{nn} < 0$ by (59). Define $E_1 \equiv \frac{n}{2\tau k}$

$\Pi_{n\tau} - \Pi_{k\tau}$ and $E_2 \equiv \Pi_{n\tau}\Pi_{cc} - \Pi_{c\tau}\Pi_{nc}$. Before we can calculate their signs, we must make an additional assumption in order to fix the sign of n^*.

Assumption 3: *At the optimum, the value of n is greater than one, i.e.,* $n^* > 1$.

Using (61) and (63) with this assumption, we get:

$$E_1 = -\sigma \ln n \left(\Pi_{k\sigma} - \frac{n^2}{2\tau^2 k\sigma}\Pi_{nn}\right) + \frac{n^2}{2\tau^2 k}\Pi_{nn}\frac{1}{1+2\tau} - \frac{q\sigma B n^{\tau+1}}{(1+2\tau)}\sqrt{\frac{r}{k}} < 0, \qquad (67)$$

where $\sigma \ln n > 0$ by Assumption 3, $\Pi_{k\sigma} - \frac{n^2}{2\tau^2 k\sigma}\Pi_{nn} > 0$ by (56), and $\Pi_{nn} < 0$ by (46). Using (61) and (62) similarly, we get:

$$E_2 = (\Pi_{cc}\Pi_{nn} - (\Pi_{nc})^2)\frac{n}{\tau}\ln n + \Pi_{cc}\left(\Pi_{nn}\frac{n}{\tau}\frac{1}{1+2\tau} - \frac{2\tau}{1+2\tau}q\sigma\sqrt{rk}Bn^{\tau}\right) > 0, \qquad (68)$$

because $\Pi_{cc}\Pi_{nn} - (\Pi_{nc})^2 > 0$ by (47), $\frac{n}{\tau}\ln n > 0$ by Assumption 3, $\Pi_{cc} < 0$, and $\Pi_{nn} < 0$ by Assumption 2. Combining (66), (67), and (68), we see that $\frac{\partial n^*}{\partial \tau} < 0$.

Behavior of $k^*(\tau)$. By applying Cramer's rule to a system analogous to (49) (with τ in place of σ), we get:

$$-\det D^2\Pi\,\frac{\partial k^*}{\partial \tau} = \Pi_{n\tau}(\Pi_{nc}\Pi_{ck} - \Pi_{nk}\Pi_{cc}) - \Pi_{c\tau}(\Pi_{nn}\Pi_{ck} - \Pi_{nc}\Pi_{nk}) + \Pi_{k\tau}(\Pi_{nn}\Pi_{cc} - (\Pi_{nc})^2).$$

To get rid of Π_{ck} and Π_{nk}, we use (53) and (54). This yields:

$$-\det D^2\Pi\,\frac{\partial k^*}{\partial \tau} = (\Pi_{nn}\Pi_{cc} - (\Pi_{nc})^2)\left(\Pi_{k\tau} - \frac{n}{2\tau k}\Pi_{n\tau}\right) > 0,$$

since $\Pi_{nn}\Pi_{cc} - (\Pi_{nc})^2 > 0$ by (47), and $\Pi_{k\tau} - \frac{n}{2\tau k}\Pi_{n\tau} = -E_1 > 0$ by (69). Therefore, $\frac{\partial k^*}{\partial \tau} > 0$. ■

REFERENCES

Acemoglu, D., P. Aghion, C. Lelarge, J. Van Reenen, and F. Zilibotti. 2006. Technology information and the decentralization of the firm. Working paper, London School of Economics.

Adams, C. 2006. Optimal team incentives with CES production. *Economics Letters* 92 (July): 143–148.

Autrey, R. 2005. Team synergy, team composition and performance measures. Working paper, Harvard Business School.

Baiman, S., and J. Demski. 1980. Economically optimal performance evaluation and control systems. *Journal of Accounting Research* 18 (Supplement): 184–220.

Baldenius, T., N. Melumad, and A. Ziv. 2002. Monitoring in multiagent organizations. *Contemporary Accounting Research* 19 (Winter): 483–511.

Bhattacharyya, S., and F. Lafontaine. 1995. Double-sided moral hazard and the nature of share contracts. *RAND Journal of Economics* 26 (Winter): 761–781.

Bushman, R., R. Indjejikian, and A. Smith. 1995. Aggregate performance measures in business unit manager compensation: The role of intrafirm interdependencies. *Journal of Accounting Research* 33 (Supplement): 101–128.

Calvo, G., and S. Wellisz. 1978. Supervision, loss of control, and the optimum size of the firm. *Journal of Political Economy* 86 (October): 943–952.

Christensen, P., and F. Sabac. 2002. Dynamic agency and accounting informativeness. Working paper, University of Alberta.

Fischer, P., and J. Hughes. 1997. Mutual monitoring and best agency contracts. *Journal of Institutional and Theoretical Economics* 153 (June): 334–355.

Guadalupe, M., and J. Wulf. 2007. The flattening firm and product market competition: The effect of trade costs and liberalization. Working paper, Columbia University.

Harris, M., and A. Raviv. 2002. Organizational design. *Management Science* 48 (July): 852–865.

Holmström, B. 1982. Moral hazard in teams. *Bell Journal of Economics* 13 (Autumn): 324–340.

Huddart, S., and P. Liang. 2005. Profit sharing and monitoring in partnerships. *Journal of Accounting and Economics* 40 (December): 153–187.

Hughes, J., L. Zhang, and J. Xie. 2005. Production externalities, congruity of aggregate signals, and optimal task assignments. *Contemporary Accounting Research* 61 (Summer): 398–408.

Ma, C. 1988. Unique implementation of incentive contracts with many agents. *Review of Economic Studies* 55 (October): 555–571.

Marschak, J., and R. Radner. 1972. *Economic Theory of Teams.* Cowles Foundation. New Haven, CT: Yale University Press.

McLaughlin, K. 1994. Individual compensation and firm performance: The economics of team incentives. Working paper, University of Chicago.

Prendergast, C. 2002. The tenuous trade-off between risk and incentives. *Journal of Political Economy* 110 (October): 1071–1102.

Qian, Y. 1994. Incentives and loss of control in an optimal hierarchy. *The Review of Economic Studies* 61 (July): 527–544

Rajan, R., and J. Wulf. 2006. The flattening firm: Evidence on the changing nature of firm hierarchies from panel data. *Review of Economics and Statistics* 88 (November): 759–773.

Ricart i Costa, J. 1988. Managerial task assignment and promotions. *Econometrica* 56 (March): 449–466.

Roberts, J. 2004. *The Modern Firm: Organizational Design for Performance and Growth.* Oxford, U.K.: Oxford University Press.

Sappington, D. 1983. Limited liability contracts between principal and agent. *Journal of Economic Theory* 29 (February): 1–21.

Simons, R. 2005. *Levers of Organization Design.* Boston, MA: Harvard Business School Press.

Varian, H. 1990. Monitoring agents with other agents. *Journal of Institutional and Theoretical Economics* 146 (March): 153–174.

Williamson, O. 1967. Hierarchical control and optimum firm size. *Journal of Political Economy* 75 (April): 123–138.

Ziv, A. 1993. Performance measures and optimal organization. *Journal of Law, Economics, and Organization* 9 (April): 30–50.

———. 2000. Information technology and optimal firm structure. *Journal of Accounting Research* 38 (Autumn): 297–328.

THE ACCOUNTING REVIEW
Vol. 83, No. 3
2008
pp. 823–853

The Use of DuPont Analysis by Market Participants

Mark T. Soliman
University of Washington

ABSTRACT: DuPont analysis, a common form of financial statement analysis, decomposes return on net operating assets into two multiplicative components: profit margin and asset turnover. These two accounting ratios measure different constructs and, accordingly, have different properties. Prior research has found that a change in asset turnover is positively related to future changes in earnings. This paper comprehensively explores the DuPont components and contributes to the literature along three dimensions. First, the paper contributes to the financial statement analysis literature and finds that the information in this accounting signal is in fact incremental to accounting signals studied in prior research in predicting future earnings. Second, it contributes to the literature on the stock market's use of accounting information by examining immediate and future equity return responses to these components by investors. Finally, it adds to the literature on analysts' processing of accounting information by again testing immediate and delayed response of analysts through contemporaneous forecast revisions as well as future forecast errors. Consistent across both groups of market participants, the results show that the information is useful as evidenced by associations between the DuPont components and stock returns as well as analyst forecast revisions. However, I find predictable future forecast errors and future abnormal returns indicating that the information processing does not appear to be complete. Taken together, the analysis indicates that the DuPont components represent an incremental and viable form of information about the operating characteristics of a firm.

Keywords: *financial statement analysis; DuPont analysis; market returns; analyst forecasts.*

JEL Classification: *M4.*

This paper is based in part on my dissertation at the Ross School of Business at the University of Michigan. I thank the members of my committee: Christopher Achen, Allan Afuah, Patricia Dechow, Russell Lundholm (Chair), and Richard Sloan. I also thank Mary Barth, Bob Bowen, Mark Bradshaw, Dan Dhaliwal (the editor), Ilia Dichev, Jeffrey Doyle, Nader Hafzalla, Michelle Hanlon, Ron Kasznik, Maureen McNichols, Sarah McVay, Venky Nagar, Madhav Rajan, Scott Richardson, Terry Shevlin, Teri Yohn, two anonymous referees, and the doctoral students at the University of Michigan for their helpful discussions and comments. This paper also has benefited greatly from the comments of workshop participants at University of California, Berkeley, University of Chicago, Columbia University, Duke University, Massachusetts Institute of Technology, Northwestern University, University of Pennsylvania, University of Southern California, and Stanford University. I gratefully acknowledge the financial support of the Ross School of Business at the University of Michigan, the William A. Paton Accounting Fund, and the Graduate School of Business at Stanford University. Excellent research assistance by Julie Suh was greatly appreciated. This paper is dedicated to the loving memory of my dear friend Nader Hafzalla.

Editor's note: This paper was accepted by Dan Dhaliwal.

Submitted February 2006
Accepted September 2007

Data Availability: *The data used in this study are publicly available from the sources indicated in the text.*

I. INTRODUCTION

In this paper I examine whether the information contained in DuPont analysis is associated with stock market returns and analyst forecasts. Prior studies document that the components from DuPont analysis, which decomposes return on net operating assets into profit margin and asset turnover, have explanatory power with respect to changes in future profitability. This paper adds to the literature by comprehensively examining investor and analyst reactions to the DuPont components along three dimensions. First, it replicates the previously documented forecasting ability and examines whether it is robust and incremental to the existence of other predictors already considered in the literature. Second, it explores the use of these components by stock market investors by looking at both contemporaneous and future returns. In contemporaneous long-window association and short-window information tests, the results indicate a positive association between the DuPont components and equity returns. But small future abnormal returns to a trading strategy indicate a possibly incomplete processing of information. Finally, it examines both current forecast revisions and future forecast errors by sell-side analysts. Although they appear to revise their forecasts of future earnings consistent with the information in these DuPont components, the revision seems to be incomplete as evidenced by predictable future forecast errors. Consistent across both groups of market participants, the contemporaneous results show that the information is useful but the future tests indicate that the information processing does not appear to be complete.

Work by Nissim and Penman (2001) provides an approach to equity valuation using the residual income framework that gives a simple direct mapping of financial ratios to equity valuation. In particular they use DuPont analysis, which decomposes a firm's return on net operating assets (*RNOA*) into profit margin (*PM*) and asset turnover (*ATO*) where $RNOA = PM \times ATO$.[1] *PM* and *ATO* are accounting signals that measure different constructs about a firm's operations. *PM* is often derived from pricing power, such as product innovation, product positioning, brand name recognition, first mover advantage, and market niches.[2] *ATO* measures asset utilization and efficiency, which generally comes from the efficient use of property, plant, and equipment; efficient inventory processes; and other forms of working capital management.[3]

There are reasons to expect competitive forces to affect these two sources of profitability differently. Large profit margins often draw new entrants into the marketplace or quick imitation of new ideas from existing rivals. The resulting competition causes high profit margins to revert to normal levels, suggesting more transitory benefits. Unlike profit margin, however, competition may be less threatening to an efficient deployment of assets. It is more difficult to imitate another firm's efficient production processes because such imitation often involves large and costly overhauls of current factories and operations.

---

[1] Specifically, $RNOA$ = Operating Income/Average Net Operating Assets, PM = Operating Income/Sales, and ATO = Sales/Average Net Operating Assets. Hereafter, *PM* and *ATO* are referred to as the "DuPont Components." Another common form is to decompose ROE = (Profit Margin × Asset Turnover × Leverage) or [*NI/Sales* × *Sales/Assets* × *Assets/Equity*]. As discussed in the "Valuation Theory and *RNOA*" section, I use *RNOA* in my analyses in order to focus on operations and thus abstract from the firm's financing decision.

[2] For example, Abercrombie and Fitch earns high margins by selling used-looking clothing that is considered trendy and is demanded by teenagers. Their strong brand is difficult to imitate and allows them to charge a premium.

[3] A good example of high abnormal *ATO* is Dell. Their business model is based on maintaining extremely low inventory and high turnover. Historically, their *ATO* has exceeded the industry median by about 1.2 turns.

Theoretical work supports these conjectures. Romer (1986) contends that knowledge is easily diffused and transferred throughout an economy, making returns based on knowledge-based assets transitory and diminishing.[4] In this setting, to the degree that the source of profit margin is derived from ideas that can be imitated by others, it is more likely to be more transitory. Moreover, Romer (1986) concludes that returns derived from capital are more persistent because of the larger frictions to the movement of capital through an economy.[5]

Work by Fairfield and Yohn (2001); Nissim and Penman 2001 and Penman and Zhang (2003) is consistent with the intuition above and shows that assert turnover is more persistent than profit margin and that changes in *ATO* are predictive of future changes in *RNOA* after controlling for *RNOA*.[6] In this paper, I begin with a replication to investigate whether this finding is robust to the inclusion of variables in existing research on fundamental analysis. I find that despite the correlation that exists between these variables, the change in *ATO* is still significant in explaining future changes in *RNOA* after controlling for the nine fundamental signals from Abarbanell and Bushee (1997; hereafter, AB) and the variables in the extended accrual decomposition of Richardson et al. (2005; hereafter, RSST).

The focus of this study, however, explores the use of the information in these components by those who have a vested interest in the prediction of future earnings. Numerous studies have explored how market participants, either stock investors or analysts, incorporate the information in earnings into their decisions. Along similar veins, this paper extends the literature on equity holders and analysts by exploring the immediate and delayed reactions to these accounting signals. These two complementary approaches guide the analyses throughout the paper: First, contemporaneous tests examine whether the information in these components is associated with current actions by market participants; and second, future tests examine whether market participants *fully* use the information.

To test whether stock market returns are associated with the DuPont components, I conduct both long-window association and short-window information tests.[7] The long-window association tests show that the DuPont components are incremental to earnings and earnings changes in explaining contemporaneous returns, and that adding them to the regression doubles the traditionally low adjusted R^2s that come with these tests (Lev 1989; Kothari 2001).[8] The short-window return tests show that changes in the DuPont components are informative to the marketplace and incremental to the earnings surprise. The two results suggest that the DuPont components not only reflect information in the marketplace (long-window return tests), but also that changes in *ATO* appear to provide information to stock market participants within a small window around the earnings announcement (short-window). Finally, a future return tests produce an annual abnormal return of about 5 percent from using the information in the change in *ATO*. Thus, although the future predictive ability of earnings is reflected in short-window announcement returns, the market does not appear to completely use the information in this variable.

---

[4] Of course, there are many ways a firm can protect abnormally high profit margins from competition, such as established marketing channels, brand name recognition, customer loyalty, unique supply sources, favorable contracts with customers or suppliers, and legal frictions such as patents and copyrights. All these factors work to promote the persistence of profit margin.

[5] Additionally, there are simple accounting reasons to expect *ATO* to be different in nature than *PM*. *ATO* is calculated as sales (a flow variable) divided by net operating assets (a stock variable); both of these line items have relatively low variance (Nissim and Penman 2001). However, *PM* is calculated by dividing operating income by sales (two flow variables), and operating income simply has higher volatility than net operating assets or sales. This also could lead to differing persistence. Thus, the very construction of these ratios and the accounting behind them also leads to expecting different time-series properties, i.e., that *PM* will be less persistent and more volatile than *ATO*. Accordingly, any shift in *ATO* is more likely to be the result of an economic occurrence rather than changes due to simple variation in the accounting numbers.

[6] For example, compare Toyota to Ford. Toyota has far more efficient and technologically advanced factories. To compete along this dimension, Ford would have to restructure and rebuild billions of dollars' worth of factories and manufacturing plants. Of course, counter examples exist to this notion. For example, Lowe's Hardware entered into Home Depot's market and successfully competed with their massive scale and efficiency.

[7] See Kothari (2001) for a summary of this large literature on the relation between returns and earnings.

[8] A supplemental finding is that adding *RNOA* along with earnings in these regressions increases the adjusted R^2 significantly. This lends support to claims that *RNOA* is a more appropriate measure when examining how accounting profitability maps into equity returns in these types of long-window regressions.

This paper also contributes to the literature on sell-side analysts' use of accounting information. If the DuPont components map into equity value, then analysts should use this information when creating forecasts and revising their priors about the future profitability of the firm. Consistent with the immediate and delayed analysis of returns above, two measures of analyst forecasts are examined. First, I use analyst forecast revisions measured as the change in the forecast of next year's earnings immediately before and after the earnings announcement. Consistent with the earnings prediction tests, I find that analysts' immediate reaction is to revise their forecasts based on changes in *ATO* after controlling for the earnings surprise. Second, I look at future forecast errors of year $t+1$ earnings and examine whether analysts fully understand the implications of the DuPont components in year t for future profitability. I find that the DuPont components have predictive power for future forecast errors, suggesting that analysts do not completely utilize the information in these components when issuing their forecasts. Taken together, the evidence suggests that although analysts appear to use the information in revising their forecasts of the following year's earnings, the updating is incomplete and analysts leave some information "on the table." This result is entirely consistent with future abnormal returns above. Thus, both equity investors and analysts appear to use some of the information available, but not all of it.

The remainder of the paper is organized as follows. The next section develops the predictions and research design. Section III describes the empirical analysis, Section IV presents the results, and Section V concludes.

II. PREDICTIONS AND RESEARCH DESIGN

Since the landmark studies of Ball and Brown (1968) and Beaver (1968), accounting researchers have worked to understand how and why earnings and security returns are associated. Beaver (1998) argues that one theoretical link between earnings and share prices is that current earnings provide information to predict future earnings. Consistent with this theme, researchers have looked for current financial statement information that aids in predicting future earnings, arguing that this should be the primary goal of fundamental analysis (Penman 1992; Lee 1999). To this end, several approaches have emerged in the literature. Lipe (1986) and Kormendi and Lipe (1987) decompose earnings into six components and find that more transitory components have a smaller relation with stock returns. Along the same lines, Fairfield et al. (1996) use the order of line items on the income statement to decompose earnings to improve predictions of *ROE* and find that line items farther down the income statement are less persistent. Ou and Penman (1989) combine a large group of financial ratios into one summary measure and estimate its association with future stock returns. Holthausen and Larcker (1992) extend this analysis and use the variables to directly predict future returns.

Lev and Thiagarajan (1993) take a different approach and identify a group of financial ratios used in practice by "experts" and examine their correspondence with contemporaneous long-window stock returns. AB (1997, 1998) extend this analysis by using the same variables and find that some of the ratios (1) predict future earnings changes, (2) are used by analysts, and (3) can predict future returns. Sloan (1996) decomposes earnings into accruals and cash flows and finds that firms with higher operating accruals tend to have lower future earnings and returns. RSST (2005) expand the definition of accruals to map out the entire balance sheet and find that investing accruals are more effective than operating accruals in predicting future earnings and returns. Although the approaches mentioned above were successful in predicting future earnings, they lack a unifying framework guiding the analysis despite the fact that most of these studies were ultimately examining equity returns.[9]

Valuation Theory and *RNOA*

The theory of finance describes stock price as a function of expected future dividends leading to the familiar dividend-discount model that sets firm value to the present value of the firms expected future dividends. An algebraically equivalent model, commonly referred to as the *residual income model*, stems from the assumption of clean-surplus accounting and expresses firm value in terms of accounting numbers.[10] The stock price can be rewritten in terms of *ROE* and reported book value plus an infinite sum of discounted residual income:

$$P_t = B_t + \sum_{i=1}^{\infty} \frac{E_t[(ROE_{t+i} - r_e)B_{t+i-1}]}{(1 + r_e)^i}$$

where:

P_t = *current stock price*;
B_t = book value at time t;
$E_t(.)$ = expectation based on information available at time t;
ROE_{t+1} = the return on book equity for period $t+1$;
$E_t(D_{t+i})$ = expected future dividends for periods $t+i$ conditional on information available at time t; and
r_e = cost of equity capital.

Work by Ohlson (1995) and Feltham and Ohlson (1995) highlights the theoretical importance of *ROE* in the implementation of valuation models in general, and in the residual income model in particular. Standard DuPont analysis decomposes *ROE* into the three multiplicative ratios of Profit Margin, Asset Turnover, and Leverage as follows:

$$ROE = [NI/Sales \times Sales/Assets \times Assets/BVEquity].$$

As the equation above shows, *ROE* can be affected by the firm's choice of capital structure, yet changes in the firm's capital structure may not be value relevant.[11] Nissim and

---

[9] Barth et al. (1999) use an Ohlson (1999) framework to justify and determine the accruals and cash flow valuation coefficients and place their analysis within a broader valuation context.

[10] The model is sometimes referred to as the Edwards-Bell-Ohlson (EBO) valuation equation. Clean surplus accounting requires that all gains and losses affecting book value be included in earnings such that changes in book value are equal to earnings minus net dividends ($b_t = b_{t-1} + NI_t - D_t$).

[11] Modigliani and Miller (1958) assert that it is the value of the operations that matters and not the financing of those assets. Penman (2001, 433–434) shows that despite the fact that *ROE* can be mechanically increased through leverage (assuming positive spread), the increase in the discount rate results in no change in equity value.

Penman (2001) algebraically rearrange *ROE* to abstract away from financial leverage and arrive at *RNOA* as follows:[12]

$$ROE = RNOA + [FLEV \times SPREAD].$$

RNOA captures the firm's operating profitability without the effects of financial leverage and is becoming commonly used in the valuation literature (e.g., Fairfield and Yohn 2001; Nissim and Penman 2001; Penman and Zhang 2003; Fairfield et al. 2003a; Richardson et al. 2006). Using this residual-income valuation framework, they algebraically derive how equity prices should map into *RNOA* and decompose it into the DuPont components. After removing leverage and any associated returns (e.g., interest income/expense), *RNOA* can be multiplicatively decomposed into *PM* and *ATO* as follows: $RNOA = PM \times ATO$ where PM = Operating Income/Sales and ATO = Sales/Net Operating Assets.

Nissim and Penman (2001) find that these two components of *RNOA* (*PM* and *ATO*) have different time-series properties. *PM* measures the firm's ability to control the costs incurred to generate sales and gives insight into the sensitivity of operating income to product price and cost structure. ΔPM measures the growth rate in operating income relative to the growth rate in sales. *ATO* captures the firm's efficiency in using operating assets to generate sales and is often interpreted as a measure of asset utilization by managers. ΔATO reflects change in the productivity of the firm's assets and ultimately measures growth in sales relative to growth in operating assets. Ultimately this measure evaluates sales growth while controlling for growth in net operating assets used to generate those sales. As discussed in the introduction, each component measures a different aspect of a firm's operations. Thus, simply using the aggregated level of *RNOA* may result in the loss of some predictive information. Two research questions immediately follow: (1) Do these two components have incremental explanatory power with respect to future earnings? (2) If so, then do market participants, such as equity analysts and stock market investors, use this information when valuing stocks? The focus of this paper is centered on the second question.

Predictions of Future *RNOA* Using DuPont Components

The goal of these first tests is to ensure that the explanatory power documented by Fairfield and Yohn (2001) and Penman and Zhang (2003) is truly incremental and robust to other earnings predictors from the literature before moving on to the larger question of whether this predictive power is used by market participants. To do this, I first control for the fundamental signals proposed by Lev and Thiagarajan (1993). These signals were later used by AB (1997) and shown to predict future earnings (as well as future returns [AB 1998]).[13] This first step is important because one of the AB variables is gross margin, which is highly correlated to *PM*. In addition, their measures of inventory, accounts receivable, and capital expenditures are all parts of *ATO*. Thus, it is quite conceivable that earlier work on DuPont analysis was simply rediscovering the same underlying constructs using slightly different measures. Second, I examine whether the information in the DuPont components is subsumed by accruals. In particular, I use the three components of total accruals in RSST (2005): working capital, noncurrent operating and financing, labeled ΔWC, ΔNCO, and

---

[12] *FLEV* is financial leverage and *SPREAD* is the difference between return of the firm's operations and borrowing costs.

[13] They include measures of inventory, accounts receivable, capital expenditures, gross margins, selling and administrative expenses, effective tax rates, earnings quality, audit quality, and labor force. See Table 3 for exact definitions.

ΔFIN. They find that ΔNCO accruals have more explanatory power of future earnings and returns than do operating accruals. This expanded definition of accruals subsumes the operating accruals definition used by Sloan (1996) and provides a more complete picture of the degree of accruals in the firm by examining the entire balance sheet.

Fairfield and Yohn (2001) were the first to address the question of future predictive power and find that ΔATO is positively associated with future changes in *RNOA*, but that levels of *PM* and *ATO* have no predictive value. This result is intuitive because the levels of *PM* and *ATO* are more informative toward a firm's operating structure or industry membership (Ge and Soliman 2007). Further, increases in *ATO* indicate that the firm's ability to generate sales from a given investment has increased and that this is an indicator of future efficiency of generating sales from assets. Thus, this type of increase in profitability tends to persist. In this section, I replicate Fairfield and Yohn's (2001) findings by applying their models to my sample. I first estimate the following regression on the *level* of the DuPont components:[14]

$$\Delta RNOA_{t+1} = \rho_0 + \rho_1 RNOA_t + \rho_2 PM_t + \rho_3 ATO_t + \rho_4 \Delta RNOA_t + \rho_5 \Delta NOA_t + \upsilon_{t+1}. \tag{1}$$

Next, I employ a changes specification similar to the one used by Fairfield and Yohn (2001) and examine the explanatory power of ΔPM and ΔATO with one notable exception. In their analysis, they do not control for changes in *RNOA* ($\Delta RNOA_t$). Thus, it is difficult to surmise from their analysis whether the explanatory power comes from ΔPM and ΔATO or whether the omitted variable, $\Delta RNOA$, is providing the information. Accordingly, I estimate the following changes regression and include $\Delta RNOA$ to determine whether changes in the components are incremental as follows:

$$\begin{aligned}\Delta RNOA_{t+1} = {} & \rho_0 + \rho_1 RNOA_t + \rho_2 \Delta PM_t + \rho_3 \Delta ATO_t + \rho_4 \Delta RNOA_t \\ & + \rho_5 \Delta NOA_t + \upsilon_{t+1}.\end{aligned} \tag{2}$$

Because pooled regressions are subject to cross-sectional correlation in the residuals, I estimate all the regressions in the paper separately for each year in the sample and then construct Fama-MacBeth t-statistics using the resulting set's time-series of annual coefficient estimates (Fama and MacBeth 1973). I also adjust for serial correlation in the annual coefficient estimates due to overlapping periods (Bernard 1987) by using the Newey and West (1987) adjustment (see Verbeek [2000] for a detailed explanation).

Thus, the first part of the analysis is primarily a replication of earlier work, but makes a small contribution in that it replicates the prior findings in a different sample and ensures that the result is robust to other control variables from the literature using slightly updated tests.[15]

---

[14] This is the same specification as model 2 of Table 2 in Fairfield and Yohn (2001). As they point out, there is no hypothesized reason to include the level of *PM* or *ATO* since they only capture a firm's operating strategy. Nonetheless, they are included in my analysis to ensure comparability with prior studies. Note that both Equations (1) and (2) are also estimated using the AB and RSST controls.

[15] That is, Fama-MacBeth regressions coupled with Newey-West adjustment for serial correlation as opposed to pooled regressions.

Use of Information by Market Participants

Stock Return Tests

In this section, I test whether stock market participants use the information contained in the components of DuPont analysis, particularly ΔATO. If the results from the profitability tests above are robust, then the natural question is whether stock market participants seem to impound the information in their decision making. The question addressed here is whether the information in the components of DuPont analysis is associated with equity returns. Stated differently, does this form of earnings decomposition (1) capture information that is relevant to market participants about the firm and its future prospects incremental to current earnings or (2) provide information to capital market participants that is associated with revised beliefs? To test these two questions, I employ long- and short-window return tests, respectively.

Long-Window Return Tests

In the spirit of Ball and Brown (1968), I examine whether there is a statistical association between information in the components of DuPont and market equity returns. This would suggest that the information in the components is correlated with information used by investors (Francis and Schipper [1999] label this as "Interpretation 4" of value relevance), and therefore captures information that affects equity valuation. Finding statistical significance is consistent with the joint hypothesis that the information in the DuPont components reflects underlying events that affect future profitability and that such events also are reflected in the market prices of securities. Unlike short-window tests that suffer from a lack of power and only capture information that surprises the market, long-window, contemporaneous, 12-month annual return tests will help ascertain whether the DuPont components contain information that is useful to the market.[16]

There is still much debate on the proper specification of the returns-earnings regression (see Kothari 2001). Kothari (1992) suggests that using an earnings-level specification reduces the bias in the coefficients because prices lead earnings. However, Easton and Harris (1991) argue that $\Delta EARN$ is also a relevant variable. Many papers simply include both (e.g., Amir and Lev 1996; Francis and Schipper 1999).[17] In addition, Ohlson (1995) claims that *ROE* (and through some algebra *RNOA*) is an important input into valuation models. If these arguments are true, then one would expect *RNOA* to have better correspondence with stock prices than earnings. Accordingly, I also add *RNOA* and $\Delta RNOA$ and estimate the following regression equation in pieces to see whether (1) *RNOA* is incremental to earnings and (2) the DuPont components are incrementally useful to earnings and *RNOA*:

$$\begin{aligned} R_t = {} & \rho_0 + \rho_1 EARN_t + \rho_2 \Delta EARN_t + \rho_3 RNOA_t + \rho_4 \Delta RNOA_t + \rho_5 PM_t \\ & + \rho_6 ATO_t + \rho_7 \Delta PM_t + \rho_8 \Delta ATO_t + \varepsilon_t \end{aligned} \tag{3}$$

---

[16] Other researchers have used 15-month windows (e.g., Francis and Schipper 1999) in order to capture the earnings announcement period. However, because I also conduct short-window return tests around the earnings announcement (described below), the 15-month window overlaps with this analysis. As a robustness check, I use a 15-month window and find similar results.

[17] As a robustness check, I estimated these regressions with one or the other of either *EARN* or $\Delta EARN$ along with the DuPont components and found similar results on my variables of interest. Brown et al. (1999) point out that using returns on the LHS can help avoid econometric issues and mitigates any intertemporal constant correlated omitted variables.

where:

R_t = stock returns are measured using compounded buy-hold market-adjusted returns (raw return minus the corresponding value-weighted return), inclusive of dividends and other distributions computed over the 12 months beginning in the first month of the firm's fiscal year and ending at the end of the fiscal year t;

$EARN_t$ = EPS_t/P_{t-1}; earnings before extraordinary items per share in year t, deflated by the market value of equity per share at the end of fiscal year $t-1$;

$\Delta EARN_t$ = $\Delta EPS_t/P_{t-1}$; earnings before extraordinary items per share in year t minus its annual earnings per share in year $t-1$, deflated by the market value of equity per share at the end of fiscal year $t-1$.

DuPont components and *RNOA* are as defined earlier.[18]

Equation (3) examines the *contemporaneous* relation between returns and earnings and whether the components of DuPont analysis are incrementally informative to stock market participants. However, as Beaver et al. (1980) point out, prices lead earnings because there is a richer information set impounded in stock prices *vis-à-vis* earnings. Thus, an alternative approach is to include future earnings as an independent variable (Warfield and Wild 1992; Fama 1990). Similar in spirit to Kothari and Sloan (1992) and Collins et al. (1994), I add future levels of *RNOA* and DuPont components to model (3) as follows:[19]

$$R_t = \rho_0 + \rho_1 EARN_t + \rho_2 \Delta EARN_t + \rho_3 RNOA_t + \rho_4 \Delta RNOA_t + \rho_5 PM_t + \rho_6 ATO_t + \rho_7 \Delta PM_t + \rho_8 \Delta ATO_t + \rho_9 RNOA_{t+1} + \rho_{10} PM_{t+1} + \rho_{11} ATO_{t+1} + \varepsilon_t. \quad (4)$$

This specification examines whether current market returns reflect *future* levels of profit margin and asset turnover in addition to future and current levels and changes of *RNOA*, *PM*, and *ATO*.

Short-Window Return Tests

Previous research has not studied the market reaction to the DuPont components and whether they are incremental to the earnings surprise.[20] Thus, holding both the expectation and the level of earnings constant, the question is whether the source of earnings is useful to market participants at the time of the earnings announcement, i.e., is the information in these DuPont variables timely?[21] Consistent with prior research, I examine the unexpected return reaction to the earnings surprise. I use changes throughout the analysis because of the short-window nature of the test. Changes in the DuPont components are added to see whether they are incrementally significant to the earnings surprise and change in *RNOA* as follows:[22]

---

[18] The DuPont components are not scaled by stock price because they are ratios. As a robustness check all DuPont variables are scaled by beginning period stock price; inferences are unchanged.

[19] For both Equations (3) and (4), I estimate the regressions with the three Fama-French factors to control for risk. Results are qualitatively similar.

[20] To fully calculate *PM* and *ATO*, market participants would need some balance sheet information and sales, in addition to operating income, at the time of the annual earnings announcement. Most firms disclose sales at the time of the earnings announcement; and Chen et al. (2002) find that about 52 percent of quarterly earnings announcements contain balance sheet information. In addition, they find the highest proportion of balance sheet disclosure occurs in the 4th quarter, corresponding to the annual earnings announcements I examine.

[21] Of course, an alternative explanation could be that the market is reacting to different information provided by the company at the time of the earnings announcement and that this information is correlated with the DuPont components (Beaver et al. 1970).

$$UR_t = \rho_0 + \rho_1 SUR_t + \rho_2 \Delta PM_t + \rho_3 \Delta ATO_t + \rho_4 \Delta RNOA_t + \varepsilon_t \tag{5}$$

where:

UR_t = stock returns are measured using compounded buy-hold market-adjusted returns (raw return minus the corresponding value-weighted return), inclusive of dividends and other distributions computed over the five-day window surrounding the earnings announcement beginning two days before and ending two days after the annual earnings announcement for fiscal year t; and

SUR_t *(Annual Earnings Surprise)* = annual I/B/E/S earnings − the most recent median forecast of annual earnings for year t deflated by the market value of equity per share at the end of fiscal year $t-1$.

Future Return Tests

The final return test is based on the growing evidence that market participants do not fully understand the time-series properties of earnings. This stream of research shows that investors sometimes do not understand the future implications of current earnings mapping into future earnings (e.g., Bernard and Thomas 1989; Sloan 1996; Doyle et al. 2003) and that a trading strategy exploiting this information can earn abnormal returns. To implement this trading strategy, I explore whether investors understand the future implications of $\Delta RNOA$ as a function of the DuPont components using the following regression:

$$\begin{aligned} R_{t+1} = {} & \rho_0 + \rho_1 \Delta RNOA_t + \rho_2 \Delta PM_t + \rho_3 \Delta ATO_t + \rho_4 RSST\ Controls \\ & + \rho_5 RNOA_t + \rho_6 PM_t + \rho_7 ATO_t + \rho_8 Fama\text{-}French\ Risk\ Factors + \varepsilon_{t+1} \end{aligned} \tag{6}$$

where:

R_{t+1} = future stock returns are measured using compounded buy-hold market-adjusted returns (raw return minus the corresponding value-weighted return), inclusive of dividends and other distributions beginning four months after the end of the fiscal year t and continuing for one year.[23]

---

[22] Because short-window return tests capture the updating of priors and represent new information, I only use changes in the DuPont components for this set of tests. Implicitly, this assumes that the market's expectations of profit margin and asset turnover are simply the prior years' expectations: a random walk model. The same reasoning holds for the testing of analyst forecast revisions below. There is the possibility that ΔATO is not a good measure of the surprise component of *ATO*. In untabulated tests, I only use innovations in the fourth quarter of *ATO* as a robustness check and find similar results.

[23] Beginning the return window four months after year-end is standard in the literature (e.g., Sloan 1996), since firms generally file Form 10-Ks within four months after the end of the fiscal year (see Alford et al. 1994). For firms that delist during the future return period, I calculate the remaining return by taking CRSP's delisting return and then reinvesting the proceeds in the value-weighted market portfolio. For firms that delist due to poor performance (delisting codes 500 and 520–584), I use a −35 percent delisting return for NYSE/AMEX firms and a −55 percent delisting return for NASDAQ firms, as recommended in Shumway (1997) and Shumway and Warther (1999). This mitigates concerns with potential survivorship biases.

If market participants fully appreciate the predictive power of the DuPont variables, then the coefficient is expected to be zero. As control variables, I include the Fama-French risk factors of book-to-market ratio (*BM*), size (*MVE*), and β (Beta) in my tests as risk factors (Fama and French 1993). In addition, I include the three accrual variables from RSST (described earlier), which also have been shown to be strong predictors of future abnormal returns. In these tests, I use rank regressions where the continuous value of the independent variable amount is replaced with its annual decile rank. These are more conservative statistical tests; the variables are scale-free and the only assumption about the regression's functional form is that the relations are monotonic (Iman and Conover 1979). To create decile ranks, all the continuous variables are sorted annually into ten equal-sized groups numbered 0 to 9 each year and then divided by 9. This allows for easy interpretation of the absolute value of the coefficient estimate on a variable as the hedge value moving from the top to the bottom decile created to optimize on the information in the variable, after controlling for the other variables in the regression (Bernard and Thomas 1989).

Analyst Forecast Tests

Because a primary task of equity analysts is to predict future earnings, it is expected that their forecasts contain all relevant information. A variety of research provides evidence consistent with differing degrees of analyst sophistication. Several studies find that analysts do not fully impound relevant accounting information into their earnings forecasts (e.g., Stober 1992). A number of studies show that analyst forecasts are inefficient in the sense that they do not fully incorporate past information available at the time of their forecasts. Lys and Sohn (1990), Klein (1990), and Abarbanell (1991) all suggest that analysts under-react to past information reflected in prices. Others have explored whether market anomalies can be traced to how information intermediaries such as equity analysts process accounting information. Abarbanell and Bernard (1992) find that although analyst forecasts under-react to earnings, the biases are not large enough to explain the post-earnings announcement drift completely. Bradshaw et al. (2001) explore whether equity analysts understand the lower persistence of accruals and find that they, like equity investors, do not fully understand accruals. Finally, AB explore whether the forecasts of equity analysts fully incorporate the information in the fundamental variables used by Lev and Thiagarajan (1993) and find that they do not.

Notwithstanding the studies above, another series of research studies indicates that the relation can go the other way, namely that equity investors can inefficiently utilize information contained in analysts' earnings forecasts. Mendenhall (1991), Walther (1997) and Elgers et al. (2001, 2003) document that investors underweigh information contained in analysts' forecasts. Thus, exploring tests of stock returns and analysts' forecasts are not perfect substitutes since neither one has been shown to subsume the other. Accordingly, whether analysts will comprehend the DuPont components in the same way equity investors do is still an open question and not immediately obvious from the results of the stock return tests.

To this end, I use two approaches to examine whether analysts fully impound the DuPont component information into their forecasts. First, I examine whether the DuPont components are associated with analyst forecast revisions of future earnings of period $t+1$ in the month immediately before and after the earnings announcement of year t. This approach examines whether the information is associated with analysts revising their prior forecasts (e.g., Barth and Hutton 2004). Significant coefficients on the DuPont components indicate that the information in these ratios is useful to analysts in revising their priors about the future prospects of the firm or is at least correlated with other information that is. Positive coefficients indicate that the analysts are revising their forecasts upward. Once again, because of the nature of the dependent variable, I use changes in this specification to see if they are incremental to the surprise and changes in *RNOA* as follows:

$$Anal\_REV_t = \rho_0 + \rho_1 SUR_t + \rho_2 \Delta PM_t + \rho_3 \Delta ATO_t + \rho_4 \Delta RNOA_t + \varepsilon_t \quad (7)$$

where:

$Anal\_REV_t$ = analyst revision of one-year-ahead earnings forecast measured as the revision to the consensus analyst forecast of year $t+1$ earnings made just after year t earnings are announced. Specifically, it is the first median I/B/E/S consensus one-year-ahead forecast of year $t+1$ earnings, minus the last median consensus of year $t+1$ earnings made directly before the announcement of year t earnings, all scaled by share price at the end of fiscal year $t-1$.

The second test on equity analysts examines whether analysts completely understand the predictive power that the DuPont components have for future profitability. In contrast to Equation (7), the following tests explore whether analysts *fully* use the information in the DuPont components by testing whether future forecast errors for period $t+1$ are predictable. This type of research design has been used by others such as Bradshaw, Richardson, and Sloan (2001). In this specification, because future forecast errors can be seen as a level or a change, I include both the levels and changes of the DuPont components as follows:

$$FE_{t+1} = \rho_1 + \rho_2 \Delta RNOA_t + \rho_3 PM_t + \rho_4 ATO_t + \rho_5 \Delta PM_t + \rho_6 \Delta ATO_t + \varepsilon_{t+1} \quad (8)$$

where:

FE_{t+1} = forecast error is the realized earnings for year $t+1$ less the median forecast earnings from the month prior to the announcement of $t+1$ earnings, scaled by the stock price at the end of the month of the earnings announcement for year t.

The analyst tests above (Equations (7) and (8)) are complementary and provide two approaches to analysts' use of information. The first tests whether analysts use the information in revising expectations and thus provides a lower bound of relevance. The future forecast error tests examine whether analysts *fully* capture all the information in the components and thus place a higher threshold of sophistication on analysts. This latter test mirrors the future abnormal return tests presented earlier. Consistently, if analysts fully impound the information in the DuPont components when creating forecasts, then the expected coefficient is zero.

III. EMPIRICAL ANALYSIS

Sample

The empirical tests employ publicly available data from three sources: I/B/E/S, CRSP, and Compustat. I/B/E/S provides analyst forecast data. I restrict the sample to the 1984–2002 period because of weak coverage by I/B/E/S in earlier years. Financial statement data are obtained from the Compustat annual database, and stock return data are obtained from the CRSP daily stock returns files. All firm-year observations with SIC codes 6000–6999 (financial companies) are excluded because the DuPont decomposition is not meaningful for these firms. This is consistent with previous studies on DuPont analysis (Nissim and Penman 2001; Fairfield and Yohn 2001) and thereby facilitates comparison among studies. Firm-year observations that are (1) not tracked by I/B/E/S, (2) have insufficient data on Compustat to compute the financial statement variables used in the tests, or (3) do not have contemporaneous and future return data on CRSP are eliminated. Requiring the sample to have I/B/E/S data biases the sample toward larger firms with analyst following.[24] In addition, all observations that do not have all the necessary information to perform my tests are removed as well as firm-year observations with negative *NOA* and operating income.[25] These criteria yield a final sample size of 38,716 firm-year observations with nonmissing analyst forecasts, financial statement, and returns data. Thus, all tests are performed on the same sample.[26]

Variable Measurement

DuPont Decomposition

RNOA is operating income before interest (Compustat item #178) divided by average net operating assets (*NOA*), where *NOA* is Operating Assets$_t$ − Operating Liabilities$_t$. Operating assets is total assets (Compustat item #6) less cash and short-term investments (Compustat item #1 and item #32). Operating liabilities is total assets (Compustat item #6), less the long- and short-term portions of debt (Compustat items #9 and #34), less book value of total common and preferred equity (Compustat items #60 and #130), less minority interest (Compustat item #38). Growth in Net Operating Assets (ΔNOA) = $(NOA_t - NOA_{t-1})/NOA_{t-1}$. *RNOA* is decomposed into the multiplicative components of *PM* (Operating Income (Compustat item #178)/Total Sales (Compustat item #12)) and *ATO* (Sales/Average *NOA*). $\Delta PM_t = (PM_t - PM_{t-1})$, and *ATO* is measured similarly.[27] Future changes in *RNOA* is $\Delta RNOA_t = RNOA_{t+1} - RNOA_t$. All financial variables are winsorized at 1 percent and 99 percent to dampen the effect of outliers on the analysis.

Control Variable Definitions

Following AB (1997), I calculate nine fundamental signals (see Table 3 for exact definitions) and consistent with RSST (2005), components of total accruals are defined as $TACC = \Delta WC + \Delta NCO + \Delta FIN$.

Finally, the control variables used to proxy for risk in the future return tests are calculated as follows:

---

[24] For example, in 2001 there were 2,707 firm-year observations with sufficient Compustat data to perform the tests of future $\Delta RNOA$ and the control variables. After requiring both I/B/E/S and CRSP, however, there are only 1,711 firm-year observations in the final sample for year 2001. It should be noted that those observations still comprise about 92 percent of the market capitalization of all Compustat firms in that year and therefore constitute a representative sample.

[25] Removing firms with negative operating income results in a loss of about 8,500 firm-year observations from my sample (about 18 percent), which is slightly less than the percentage Hayn (1995) reports because operating income is further up the income statement and does not include special items. This choice was made to (1) facilitate comparison with other studies (Fairfield and Yohn 2001), (2) remove firm-year observations where earnings do not map clearly into equity valuation (Burgstahler and Dichev 1997; Collins et al. 1999), and (3) allow for a clear ranking of performance. As an empirical matter, all the analysis was performed on the entire sample including loss firms with extremely similar results.

[26] As mentioned before, using delisting returns helps mitigate concerns of survivorship bias caused by requiring one-year-ahead future returns.

[27] As a robustness check, I also measure ΔPM and ΔATO in identical fashion as FY2001 (e.g., $\Delta PM_t = (PM_t - PM_{t-1})*ATO_{t-1}$). Inferences are not changed.

BM = book value of equity (Compustat item #60)/Market Value of Equity (shares × price − item #25 × item #199);

MVE (*Size of the Firm*) = Log(Market Value of Equity); and

BETA (β) = β for firm i for fiscal year t is estimated by a market model regression. The regression is run using weekly returns for a period of two years ending at the end of the fiscal year from which I obtain the data to compute each of the financial ratios.

IV. RESULTS

Table 1 presents descriptive statistics for the variables in the analysis. Most of the values are consistent with prior literature. *RNOA* and *PM* are slightly higher than the values reported in Nissim and Penman (2001) because (1) loss firms are excluded in this analysis and (2) I/B/E/S coverage is required in the sample, and analysts generally follow larger and more profitable firms (McNichols and O'Brien 1997). In addition, the distributions of both *PM* and *ATO* show that the means are larger than the medians, indicating positive skewness. Although not reflected in the 75th percentile, large outlier values of these ratios are present in the data. Table 2 presents Pearson and Spearman rank correlations between the variables of interest. Looking at the Spearman correlations, there is a strong negative correlation between *PM* and *ATO* at −0.364, consistent with Nissim and Penman (2001).

Future Δ*RNOA*

The prediction of future changes in *RNOA* ($\Delta RNOA_{t+1}$) is presented in Table 3. Panel A presents the results examining the levels of *PM* and *ATO*. Model 1 shows the DuPont components with changes in *NOA* and *RNOA* as control variables.[28] Consistent with prior literature, neither *PM* nor *ATO* predicts future changes in *RNOA*. In models 2, 3, and 4, I control for earnings predictors from prior literature. First, I examine the nine fundamental signals used by AB (1997) in model 2. Although the coefficients are not reported, most of the AB signals are consistent in sign and magnitude with AB (1997) with the exception of *AB_GM*, which is significant and positive in my analysis and negative in theirs.[29] In model 3, I control for the expanded definition of accruals from RSST (2005), and finally model 4 controls for all the variables simultaneously. In general, the RSST variables are still significant and consistent with prior work (i.e., Δ*FIN* is no longer significant). Interestingly, in the final specification when the AB variables are included along with the RSST variables, *ATO* is positive and significant although it is not significant in any of the other specifications or in prior literature (Fairfield and Yohn 2001; Soliman 2004). In unreported robustness checks, the significance appears to occur when adding the *CAPEX* variable, indicating that prior studies may have been misspecified without controlling for increases in capital expenditures above the industry average.

Finally, Panel B of Table 3 presents the changes specification used in prior studies. Consistent with prior literature (Fairfield and Yohn 2001; Penman and Zhang 2003), model 1 shows that only Δ*ATO* is positive and significant in predicting future changes in *RNOA*.[30] In models 2, 3, and 4, I show that this significance survives the inclusion of all the earnings

---

[28] This is done primarily to mimic model 2 in Fairfield and Yohn (2001). Fairfield et al. (2003b) argue that Δ*NOA* is a proxy for growth find and that it subsumes the operating accrual anomaly proposed by Sloan (1996).

[29] Many design differences can explain this inconsistency between this paper and Abarbanell and Bushee (1997): (1) a different sample period, (2) use of Δ*RNOA* in this analysis versus Δ*EPS*, or (3) the inclusion of *PM* in the regression which is highly correlated with Gross Margin (*AB_GM*).

[30] With the notable exception that prior studies did not include (*RNOA* as a control variable.

TABLE 1
Descriptive Statistics

	Mean	Std. Dev.	25%	Median	75%
NOA_t	1,122	2,997	44.4	160.1	691.2
$RNOA_t$	0.279	0.354	0.113	0.181	0.300
PM_t	0.115	0.082	0.056	0.096	0.153
ATO_t	2.85	2.81	1.34	2.11	3.29
$\Delta RNOA_t$	0.000	0.431	−0.048	−0.002	0.038
ΔPM_t	0.000	0.058	−0.017	0.000	0.016
ΔATO_t	−0.009	0.150	−0.154	−0.022	0.140
$Anal\_REV_t$	−0.011	0.024	−0.003	0.000	0.001
SUR_t	−0.005	0.022	−0.004	0.000	0.002
R_t	0.066	0.608	−0.278	−0.033	0.244
$EARN_t$	0.128	0.794	0.037	0.076	0.137
$\Delta EARN_t$	0.009	0.213	−0.000	0.000	0.002

The sample size is 38,716 firm-year observations from 1984–2002. Variables are winsorized at the 1 percent and 99 percent levels.

Variable definitions:

NOA (Net Operating Assets) = Operating Assets minus Operating Liabilities; Operating Assets is calculated as total assets (item #6) less cash and short-term investments (item #1) and (item #32); Operating Liabilities is calculated as total assets less total debt (items #9 and #34), less book value of total common and preferred equity (items #60 and #130), less minority interest (item #38);

PM_t *(Profit Margin)* = Operating Income$_t$ (item #178)/Total Sales$_t$ (item #12);

ATO_t *(Asset Turnover)* = Sales$_t$/Average NOA_t ((NOA_t + NOA_{t-1})/2);

$RNOA_t$ *(Return on Net Operating Assets)* = $PM_t \times ATO_t$;

$\Delta RNOA_t$ = change in *RNOA* from year t–1 through t ($RNOA_t - RNOA_{t-1}$);

ΔPM_t = change in *Profit Margin* ($PM_t - PM_{t-1}$);

ΔATO_t = change *Asset Turnover* ($ATO_t - ATO_{t-1}$);

$Anal\_REV_t$ *(Analyst Revision)* = one-year-ahead earnings forecasts measured as the revision to the consensus analyst forecast of year $t+1$ earnings made just after year t earnings are announced; specifically, it is the first median I/B/E/S consensus one-year-ahead forecast of year $t+1$ earnings minus the last median consensus of year $t+1$ earnings made directly before the announcement of year t earnings, all scaled by share price at the end of fiscal year $t-1$;

SUR_t *(Annual Earnings Surprise)* = annual I/B/E/S earnings minus the most recent median forecast of annual earnings for year t deflated by the market value of equity per share at the end of fiscal year $t-1$;

R_t = stock returns are measured using compounded buy-hold market-adjusted returns (raw return minus the corresponding value-weighted return), inclusive of dividends and other distributions computed over the 12 months beginning in the first month of the firm's fiscal year and ending at the end of the fiscal year t;

$EARN_t$ = EPS_t/P_{t-1}; the firm's earnings before extraordinary items per share in year t deflated by the market value of equity per share at the end of fiscal year $t-1$; and

$\Delta EARN_t$ = $\Delta EPS_t/P_{t-1}$; the firm's earnings before extraordinary items per share in year t minus its annual earnings per share in year t–1 deflated by the market value of equity per share at the end of fiscal year t–1.

predictors that exist in the current literature, particularly those of AB and RSST. Thus, despite the obvious collinearity that may exist between the DuPont components and the control variables, the predicative power of the ΔATO remains statistically significant. This

TABLE 2

Correlation Matrix: Pearson (above Diagonal) and Spearman (below Diagonal) (p-values shown in parentheses below correlations)

	$\Delta RNOA_{t+1}$	$\Delta RNOA_t$	PM_t	ATO_t	$RNOA_t$	$Anal\_Rev_t$	SUR_t	ΔPM_t	ΔATO_t	R_t
$\Delta RNOA_{t+1}$	—	−0.194 (0.0001)	−0.137 (0.0001)	−0.411 (0.0001)	−0.543 (0.0001)	0.071 (0.0001)	0.121 (0.0001)	−0.081 (0.0001)	−0.002 (0.676)	0.002 (0.757)
$\Delta RNOA_t$	0.0419 (0.0001)	—	0.111 (0.0001)	0.209 (0.0001)	0.319 (0.0001)	0.033 (0.0001)	0.025 (0.0001)	0.350 (0.0001)	0.292 (0.0001)	0.155 (0.0001)
PM_t	−0.151 (0.0001)	0.196 (0.0001)	—	−0.194 (0.0001)	0.373 (0.0001)	0.028 (0.0001)	0.106 (0.0001)	0.333 (0.0001)	0.011 (0.052)	0.125 (0.0001)
ATO_t	−0.157 (0.0001)	0.122 (0.0001)	−0.364 (0.0001)	—	0.655 (0.0001)	0.012 (0.025)	0.033 (0.0001)	0.026 (0.0001)	0.148 (0.0001)	0.108 (0.0001)
$RNOA_t$	−0.273 (0.0001)	0.298 (0.0001)	0.502 (0.0001)	0.536 (0.0001)	—	0.016 (0.002)	0.086 (0.0001)	0.227 (0.0001)	0.130 (0.0001)	0.176 (0.0001)
$Anal\_Rev_t$	0.187 (0.0001)	−0.189 (0.0001)	0.0731 (0.0001)	0.0627 (0.0001)	0.1244 (0.0001)	—	0.046 (0.0001)	0.063 (0.0001)	0.048 (0.0001)	0.106 (0.0001)
SUR_t	0.2494 (0.0001)	0.0693 (0.0001)	0.0409 (0.0001)	0.0416 (0.0001)	0.0747 (0.0001)	0.084 (0.0001)	—	0.066 (0.0001)	0.039 (0.0001)	0.068 (0.0001)
ΔPM_t	−0.018 (0.0001)	0.728 (0.0001)	0.2958 (0.0001)	0.055 (0.0001)	0.311 (0.0001)	0.179 (0.0001)	0.060 (0.0001)	—	0.109 (0.0001)	0.235 (0.0001)
ΔATO_t	0.1065 (0.0001)	0.6043 (0.0001)	0.0148 (0.0001)	0.124 (0.0001)	0.112 (0.0001)	0.117 (0.0001)	0.050 (0.0001)	0.179 (0.0001)	—	0.203 (0.0001)
R_t	0.0975 (0.0001)	0.3581 (0.0001)	0.177 (0.0001)	0.090 (0.0001)	0.247 (0.0001)	0.238 (0.0001)	0.090 (0.0001)	0.338 (0.0001)	0.215 (0.0001)	—

(continued on next page)

TABLE 2 (continued)

The sample size is 38,716 firm-year observations from 1984–2002. Variables are winsorized at the 1 percent and 99 percent levels.

Variable definitions:

NOA (*Net Operating Assets*) = Operating Assets minus Operating Liabilities; Operating Assets is calculated as total assets (item #6) less cash and short-term investments (item #1) and (item #32); Operating Liabilities is calculated as total assets less total debt (items #9 and #34), less book value of total common and preferred equity (items #60 and #130), less minority interest (item #38);

PM_t (*Profit Margin*) = Operating Income_t (item #178)/Total Sales_t (item #12);

ATO_t (*Asset Turnover*) = Sales_t/Average NOA_t $((NOA_t + NOA_{t-1})/2)$;

$RNOA_t$ (*Return on Net Operating Assets*) = $PM_t \times ATO_t$;

$\Delta RNOA_t$ = change in *RNOA* from year $t-1$ through t $(RNOA_t - RNOA_{t-1})$;

ΔPM_t = change in *Profit Margin* $(PM_t - PM_{t-1})$;

ΔATO_t = change *Asset Turnover* $(ATO_t - ATO_{t-1})$;

$Anal\_REV_t$ (*Analyst Revision*) = one-year-ahead earnings forecasts measured as the revision to the consensus analyst forecast of year $t+1$ earnings made just after year t earnings are announced; specifically, it is the first median I/B/E/S consensus one-year-ahead forecast of year $t+1$ earnings minus the last median consensus of year $t+1$ earnings made directly before the announcement of year t earnings, all scaled by share price at the end of fiscal year $t-1$;

SUR_t (*Annual Earnings Surprise*) = annual I/B/E/S earnings minus the most recent median forecast of annual earnings for year t deflated by the market value of equity per share at the end of fiscal year $t-1$; and

R_t = stock returns are measured using compounded buy-hold market-adjusted returns (raw return minus the corresponding value-weighted return), inclusive of dividends and other distributions computed over the 12 months beginning in the first month of the firm's fiscal year and ending at the end of the fiscal year t.

TABLE 3
Time-Series Means and t-Statistics for Coefficients from Annual Cross-Sectional Regressions of Future One-Year-Ahead Change in *RNOA* ($\Delta RNOA_{t+1}$) on the DuPont Components

Panel A: Levels of DuPont Components

$$\Delta RNOA_{t+1} = \rho_0 + \rho_1 RNOA_t + \rho_2 PM_t + \rho_3 ATO_t + \rho_4 \Delta RNOA_t + \rho_5 \Delta NOA_t + RSST\ Controls + AB\ Controls + \upsilon_{t+1}$$

Independent Variables	Model 1	Model 2	Model 3	Model 4
Intercept	0.038	0.066	0.045	0.072
	(2.97)	(3.71)	(3.28)	(4.36)
$RNOA_t$	−0.263	−0.265	−0.365	−0.257
	(−4.53)	(−5.58)	(−4.63)	(−5.25)
PM_t	0.220	0.107	0.321	0.121
	(1.17)	(1.18)	(1.26)	(1.16)
ATO_t	−0.005	−0.006	0.002	0.006
	(−1.37)	(−1.30)	(−1.53)	(2.14)
$\Delta RNOA_t$	−0.045	−0.066	−0.061	−0.061
	(−3.31)	(−2.66)	(−3.25)	(−2.81)
ΔNOA_t	−0.055			
	(−5.92)			
ΔWC_t			−0.359	−0.266
			(−7.08)	(−3.01)
ΔNCO_t			−0.212	−0.174
			(−10.53)	(−4.91)
ΔFIN_t			−0.134	−0.052
			(−5.40)	(−0.92)
AB Controls	Not Included	Included	Not Included	Included
Adjusted R^2	16.8%	17.5%	16.8%	19.8%

Panel B: Changes in DuPont Components

$$\Delta RNOA_{t+1} = \rho_0 + \rho_1 RNOA_t + \rho_2 \Delta PM_t + \rho_3 \Delta ATO_t + \rho_4 \Delta RNOA_t + \rho_5 \Delta NOA_t + RSST\ Controls + AB\ Controls + v_{t+1}$$

Independent Variables	Model 1	Model 2	Model 3	Model 4
Intercept	0.044	0.057	0.049	0.065
	(5.45)	(4.88)	(5.47)	(5.80)
$RNOA_t$	−0.277	−0.285	−0.279	−0.277
	(−6.07)	(−5.75)	(−5.95)	(−5.32)
ΔPM_t	0.074	0.155	0.069	0.338
	(1.78)	(1.65)	(1.83)	(1.26)
ΔATO_t	0.017	0.016	0.019	0.011
	(4.29)	(2.95)	(4.44)	(2.05)
$\Delta RNOA_t$	−0.078	−0.090	−0.076	−0.094
	(−2.85)	(−3.51)	(−2.93)	(−3.31)
ΔNOA_t	−0.062			
	(−9.92)			

(*continued on next page*)

TABLE 3 (continued)

Independent Variables	Model 1	Model 2	Model 3	Model 4
ΔWC_t			−0.321 (−4.57)	−0.248 (−3.31)
ΔNCO_t			−0.176 (−8.29)	−0.149 (−4.93)
ΔFIN_t			−0.098 (−3.42)	−0.044 (−0.74)
AB Controls	Not Included	Included	Not Included	Included
Adjusted R^2	16.9%	17.5%	16.8%	21.3%

The sample consists of 38,716 firm-year observations from 1984–2002. Regression results are based on 19 annual regressions using the Fama-MacBeth approach of averaging coefficients and calculating the t-statistics as follows: $\hat{\rho} \Big/ \frac{StdDev\hat{\rho}}{\sqrt{n}}$ where $\hat{\rho}$ is the simple average of the annual coefficients, n is the number of annual regressions and $StdDev\hat{\rho}$ is the standard deviation of the annual coefficients.

NOA (*Net Operating Assets*) = Operating Assets minus Operating Liabilities; Operating Assets is calculated as total assets (item #6) less cash and short-term investments (item #1) and (item #32); Operating Liabilities is calculated as total assets less total debt (items #9 and #34), less book value of total common and preferred equity (items #60 and #130), less minority interest (item #38).

ΔNOA (*Growth in Net Operating Assets*) = $(NOA_t - NOA_{t-1})/NOA_{t-1}$;

PM_t = Operating Income$_t$ (item #178)/Total Sales$_t$ (item #12);

ATO_t = Sales$_t$/Average NOA_t $((NOA_t + NOA_{t-1})/2)$;

$RNOA_t = PM_t \times ATO_t$;

$\Delta RNOA_{t+1} = RNOA_{t+1} - RNOA_t$;

$\Delta RNOA_t = RNOA_t - RNOA_{t-1}$;

$\Delta PM_t = PM_t - PM_{t-1}$;

$\Delta ATO_t = ATO_t - ATO_{t-1}$;

$\Delta WC_t = WC_t - WC_{t-1}$; *WC* is calculated as Current Operating Assets (COA) − Current Operating Liabilities (COL), COA = current assets (Compustat Item #4) − cash and Short-Term Investments (STI) (Compustat Item #1), and *COL* = current liabilities (Compustat Item #5) − debt in current liabilities (Compustat Item #34);

ΔNCO_t = change in net noncurrent operating assets is defined as $NCO_t - NCO_{t-1}$; *NCO* is calculated as Noncurrent Operating Assets (NCOA) − Noncurrent Operating Liabilities (NCOL), *NCOA* = total assets (Compustat Item #6) − current assets (Compustat Item #4) − investments and advances (Compustat Item #32), and *NCOL* = total liabilities (Compustat Item #181) − current liabilities (Compustat Item #5) − long-term debt (Compustat Item #9); and

ΔFIN_t = change in net financial assets is defined as $FIN_t - FIN_{t-1}$ and *FIN* = Financial Assets (FINA) − Financial Liabilities (FINL), FINA = Short-Term Investments (STI) (Compustat Item #193) + Long-Term Investments (LTI) (Compustat Item #32), and FINL = long-term debt (Compustat Item #9) + debt in current liabilities (Compustat Item #34) + preferred stock (Compustat Item #130).

AB Controls are comprised of the following nine variables

AB_INV = ΔInventory (#3) − ΔSales;

AB_AR = ΔAccounts Receivable (#2) − ΔSales;

AB_CAPEX = ΔIndustry Capex − ΔFirm Capex (#30);

AB_GM = ΔSales − ΔGross Margin (#12-41);

AB_ETR = Effective Tax Rate $\left[\left(\frac{1}{3}\sum_{\tau=1}^{3} ETR_{t-\tau}\right) - ETR_t\right] \times \Delta EARN_t$ where $ETR_t = \frac{TaxExpense(\#16)_t}{EBT(\#170 + \#65)_t}$;

AB_EQ = 0 for LIFO, 1 for FIFO or other;

AB_AQ = 0 for Unqualified, 1 for Qualified or other;

AB_SandA = ΔSelling and Admin Expenses (#189) − ΔSales; and

AB_LF = (Past Sales/Past Employees (#29) − Sales/Employees)/(Past Sales/Past Employees).

supports the prior literature findings and is consistent with the notion that changes in asset turnover bring new information in predicting future changes in profitability and is consistent with the economic intuition that innovations in *ATO* reflect increases in the efficiency of asset usage in generating revenues. This change in firm operational efficiency does not appear to be captured by any of the accounting signals from the prior literature included in the analysis.

Stock Return Tests

Table 4 presents the results from the long-window association tests of Equation (3). The first regression model shows that *ΔEARN* is significant in explaining returns, whereas

TABLE 4

Time-Series Means and t-Statistics for Coefficients from Annual Cross-Sectional Regressions of Contemporaneous Returns on DuPont Components

OLS Regressions for Market-Adjusted Returns

$$R_t = \rho_0 + \rho_1 EARN_t + \rho_2 \Delta EARN_t + \rho_3 RNOA_t + \rho_4 RNOA_t + \rho_5 PM_t + \rho_6 ATO_t + \rho_7 \Delta PM_t + \rho_8 \Delta ATO_t + \varepsilon_t$$

	Intercept	*EARN*	*ΔEARN*	$RNOA_t$	$\Delta RNOA_t$	PM_t	ATO_t	ΔPM_t	ΔATO_t	Adj. R^2
Model 1	0.0443	0.224	2.795							4.82%
	(1.85)	(1.43)	(2.44)							(2.65)
Model 2	−0.076	0.334	2.826	0.381	0.668					11.46%
	(−3.27)	(1.84)	(2.77)	(9.96)	(6.70)					(6.45)
Model 3	−0.129	0.308	2.885	0.284	0.067	0.496	0.006			12.48%
	(−4.73)	(1.70)	(2.76)	(6.84)	(6.69)	(2.99)	(2.36)			(7.17)
Model 4	−0.123	0.313	2.839	0.278	−0.003	0.495	0.004	0.122	0.089	14.00%
	(−4.78)	(1.63)	(2.68)	(7.04)	(−0.09)	(3.00)	(1.30)	(1.93)	(6.45)	(6.85)

The sample consists of 38,716 firm-year observations from 1984–2002. Regression results are based on 19 annual regressions using the Fama-MacBeth approach of averaging coefficients and calculating the t-statistics as follows: $\hat{\rho} \Big/ \frac{StdDev\rho}{\sqrt{n}}$ where $\hat{\rho}$ is the simple average of the annual coefficients, n is the number of annual regressions and $StdDev\rho$ is the standard deviation of the annual coefficients.

R_t = stock returns are measured using compounded buy-hold market-adjusted returns (raw return minus the corresponding value-weighted return), inclusive of dividends and other distributions computed over the 12 months beginning in the first month of the firm's fiscal year and ending at the end of the fiscal year t;

$EARN_t$ = EPS_t / P_{t-1}; the firm's earnings before extraordinary items per share in year t deflated by the market value of equity per share at the end of fiscal year $t-1$;

$\Delta EARN_t$ = $\Delta EPS_t / P_{t-1}$; the firm's earnings per share before extraordinary items per share in year t minus its annual earnings in year $t-1$ deflated by the market value of equity per share at the end of fiscal year $t-1$;

NOA (*Net Operating Assets*) = Operating Assets − Operating Liabilities; Operating Assets is calculated as total assets (item #6) less cash and short-term investments (item #1) and (item #32); Operating Liabilities is calculated as total assets less total debt (items #9 and #34), less book value of total common and preferred equity (items #60 and #130), less minority interest (item #38);

PM_t = Operating Income_t (item #178)/Total Sales_t (item #12);

ATO_t = Sales_t/Average NOA_t $((NOA_t + NOA_{t-1})/2)$;

$RNOA_t$ = $PM_t \times ATO_t$;

$\Delta RNOA_t$ = $RNOA_t - RNOA_{t-1}$;

ΔPM_t = $PM_t - PM_{t-1}$; and

ΔATO_t = $ATO_t - ATO_{t-1}$.

EARN is not. Adjusted R^2s are at levels consistent with prior literature—about 4 percent.[31] Model 2 changes the analysis slightly to focus on different measures of profitability. Instead of just using *EARN* and $\Delta EARN$, *RNOA* and $\Delta RNOA$ are added to the analysis. Many in the valuation literature are using *RNOA* as a better measure of economic performance (e.g., Penman and Zhang 2003; Fairfield and Yohn 2001; Nissim and Penman 2001; Richardson et al. 2006; Fairfield et al. 2003a). Model 2 bears out this choice. The results show that simply adding *RNOA* and $\Delta RNOA$ increases the adjusted R^2 from 4.8 percent to 11.5 percent. An (untabulated) Vuong (1989) test and F-test confirms that this increase is statistically significant. Thus, these variables are not only incrementally significant to *EARN*, but also dramatically increase the traditionally low R^2s that come with this type of analysis (Lev 1989).

When *PM* and *ATO* are added to the regression in model 3, both components are significant and positive with t-statistics of 2.99 and 2.36, respectively. This is somewhat surprising given that the levels of these variables did not predict future changes in profitability in table 3.[32] Further, it is interesting that *RNOA* is still significant in models 3 and 4, especially given that it is simply $PM \times ATO$, both of which are already in the regression equation. Thus, not only is *PM* and *ATO* incremental to *RNOA*, but *RNOA* appears to capture more information than is contained in its parts as well as in *EARN*. Finally, when DuPont component changes are added in the final specification (model 4), we again find that ΔATO is significant and positive, consistent with the incremental explanatory power documented in Panel B of Table 3, but that $\Delta RNOA$ is not. Apparently, much of the strong predictive power of $\Delta RNOA$ in models 2 and 3 is driven by ΔATO and not ΔPM. Table 4 highlights several findings. First, ΔATO is the only change component that is significantly associated with contemporaneous returns consistent with the earnings prediction results. Second, the explanatory power of the return regression increases significantly with the addition of *RNOA* to *EARN*. And finally, despite the lack of earnings predictive ability, both *PM* and *ATO* are significant in explaining contemporaneous returns. This indicates that the DuPont components are incremental to earnings in capturing the information relevant to investors in pricing securities with explanatory power as high as 14 percent when examining *RNOA* and the DuPont components.

Table 5 presents the results of the tests of Equation (4)—whether future levels of *PM* and *ATO* in year $t+1$ are associated with contemporaneous returns in year t. The coefficient ρ_{11} on ATO_{t+1} is significant with t-statistic of 3.44. The results indicate that this future component is significant and incremental to current and future earnings. Once again, this points to the importance of asset turnover as a useful and informative ratio on the economics of the firm. Also note that, consistent with Table 4, *RNOA* remains significant even after adding the current and future levels of its components, *PM* and *ATO*. Not only are market returns associated with current changes in *ATO*, but market participants appear to positively price future changes in *ATO* and the efficient deployment of capital it signifies.

Table 6 provides the results of tests of Equation (5). Recall from Section II that only changes in the DuPont components are included in these return tests because of the nature of the short-window tests. In model 1, the coefficient on surprise (*SUR*) is positive and

---

[31] Prior literature has found that in the same regression *EARN* is significant, whereas $\Delta EARN$ is not (e.g., Amir and Lev 1986). In unreported results, when I estimate this regression equation using the entire sample (including loss firms), I find that both components are statistically significant and positive. Once again, different sample composition and sample periods may explain this difference.

[32] In unreported rank regressions where the scale is consistent across both variables, *PM* has a larger coefficient than *ATO* indicating that *PM* has a somewhat larger relative importance than *ATO* in explaining contemporaneous returns.

TABLE 5
Time-Series Means and t-Statistics for Coefficients from Annual Cross-Sectional Regressions of Contemporaneous Returns on Future DuPont Components

OLS Regressions for Market-Adjusted Returns

$$R_t = \rho_0 + \rho_1 EARN_t + \rho_2 \Delta EARN_t + \rho_3 RNOA_t + \rho_4 \Delta RNOA_t + \rho_5 PM_t + \rho_6 ATO_t + \rho_7 \Delta PM_t + \rho_8 \Delta ATO_t + \rho_9 RNOA_{t+1} + \rho_{10} PM_{t+1} + \rho_{11} ATO_{t+1} + \varepsilon_t$$

	Intercept	$EARN$	$\Delta EARN$	$RNOA_t$	$\Delta RNAO_t$	PM_t	ATO_t	ΔPM_t	ΔATO_t	$RNOA_{t+1}$	PM_{t+1}	ATO_{t+1}	Adj. R^2
Mean Coef.	−0.123	0.314	2.890	0.217	0.001	0.472	0.004	0.117	0.088	0.080			14.61%
t-statistic	(−4.68)	(1.64)	(2.74)	(5.07)	(0.02)	(2.67)	(1.10)	(1.81)	(6.55)	(3.51)			(7.39)
Mean Coef.	−0.140	0.312	2.863	0.289	−0.005	−0.086	−0.008	0.125	0.087	−0.007	0.724	0.017	16.04%
t-statistic	(−5.42)	(1.64)	(2.79)	(7.58)	(−0.13)	(−0.32)	(−0.68)	(1.89)	(5.88)	(−0.43)	(1.83)	(3.44)	(7.95)

The sample consists of 38,716 firm-year observations from 1984–2002. Regression results are based on 19 annual regressions using the Fama-MacBeth approach of averaging coefficients and calculating the t-statistics as follows: $\hat{\rho} \Big/ \frac{StdDev\rho}{\sqrt{n}}$ where $\hat{\rho}$ is the simple average of the annual coefficients, n is the number of annual regressions and $StdDev\rho$ is the standard deviation of the annual coefficients.

R_t = stock returns are measured using compounded buy-hold market-adjusted returns (raw return minus the corresponding value-weighted return), inclusive of dividends and other distributions computed over the 12 months beginning in the first month of the firm's fiscal year and ending at the end of the fiscal year t;

$EARN_t$ = EPS_t/P_{t-1}; the firm's earnings before extraordinary items per share in year t deflated by the market value of equity per share at the end of fiscal year $t-1$;

$\Delta EARN_t$ = $\Delta EPS_t/P_{t-1}$; the firm's earnings per share before extraordinary items per share in year t minus its annual earnings in year t–1 deflated by the market value of equity per share at the end of fiscal year $t-1$;

NOA (*Net Operating Assets*) = Operating Assets – Operating Liabilities; Operating Assets is calculated as total assets (item #6) less cash and short-term investments (item #1) and (item #32); Operating Liabilities is calculated as total assets less total debt (items #9 and #34), less book value of total common and preferred equity (items #60 and #130), less minority interest (item #38);

PM_t = Operating Income$_t$ (item #178)/Total Sales$_t$ (item # 12);

ATO_t = Sales$_t$/Average NOA_t $((NOA_t + NOA_{t-1})/2)$;

$RNOA_t$ = $PM_t \times ATO_t$;

$\Delta RNOA_t$ = $RNOA_t - RNOA_{t-1}$;

ΔPM_t = $PM_t - PM_{t-1}$; and

ΔATO_t = $ATO_t - ATO_{t-1}$.

TABLE 6
Time-Series Means and t-Statistics for Coefficients from Annual Cross-Sectional Regressions of Short-Window Unexpected Returns on the DuPont Components

$$UR_t = \rho_0 + \rho_1 SUR_t + \rho_2 \Delta PM_t + \rho_3 \Delta ATO_t + \rho_4 \Delta RNOA_t + \varepsilon_t$$

	Intercept	SUR_t	ΔPM_t	ΔATO_t	$\Delta RNOA_t$	Adj. R^2
Model 1	0.006	0.119				0.19%
	(5.07)	(4.32)				(2.54)
Model 2	0.007	0.145	−0.001	0.003		0.21%
	(5.17)	(4.53)	(−0.72)	(3.86)		(1.07)
Model 3	0.006	0.144	−0.004	0.002	0.003	0.24%
	(5.21)	(4.59)	(−1.55)	(2.01)	(1.40)	(1.08)

The sample consists of 38,716 firm-year observations from 1984–2002. Regression results are based on 19 annual regressions using the Fama-MacBeth approach of averaging coefficients and calculating the t-statistics as follows: $\hat{\rho} \Big/ \frac{StdDev\hat{\rho}}{\sqrt{n}}$ where $\hat{\rho}$ is the simple average of the annual coefficients, n is the number of annual regressions and $StdDev\hat{\rho}$ is the standard deviation of the annual coefficients.

UR_t = stock returns are measured using compounded buy-hold market-adjusted returns (raw return minus the corresponding value-weighted return), inclusive of dividends and other distributions computed over the five-day window surrounding the earnings announcement beginning two days before and ending two days after the annual earnings announcement for fiscal year t;

SUR_t (*Annual Earnings Surprise*) = annual I/B/E/S earnings − the most recent median forecast of annual earnings for year t;

NOA (*Net Operating Assets*) = Operating Assets − Operating Liabilities; Operating Assets is calculated as total assets (item #6) less cash and short-term investments (item #1) and (item #32); Operating Liabilities is calculated as total assets less total debt (items #9 and #34), less book value of total common and preferred equity (items #60 and #130), less minority interest (item #38);

PM_t = Operating Income$_t$ (item #178)/Total Sales$_t$ (item #12);

ATO_t = Sales$_t$/Average NOA_t $((NOA_t + NOA_{t-1})/2)$;

$RNOA_t$ = $PM_t \times ATO_t$;

$\Delta RNOA_t$ = $RNOA_t - \mathrm{RNOA}_{t-1}$;

ΔPM_t = $PM_t - PM_{t-1}$; and

ΔATO_t = $ATO_t - ATO_{t-1}$.

significant, as would be expected. In the second model, only ΔATO is significant and positive, which is consistent with its implications for future earnings from Panel B of Table 3. Positive changes in *ATO* are associated with positive abnormal returns around the earnings announcement, and the result holds even after controlling for the change of profitability ($\Delta RNOA$) above and beyond the surprise in the final regression.

In my final return test, a trading strategy is explored that exploits the information in these variables. Table 7 presents the results. Model 1 indicates that there is a correspondence between ΔATO and future abnormal stock returns. The abnormal future return on such a simple signal is surprising, but is consistent with the future forecast error tests presented below. Market participants do not appear to fully impound the future predictive power of ΔATO for future changes in profitability found in Table 3. Additionally, the chance that ΔATO is simply measuring risk is also minimized by the low increase in adjusted R^2 in Table 4 when adding the DuPont components. As mentioned earlier, the coefficients are decile-ranked in order to facilitate interpretation. The coefficient on ΔATO indicates that a trading strategy optimizing the information in ΔATO would result in a hedge return of

TABLE 7
Time-Series Means and t-Statistics for Coefficients from Annual Cross-Sectional Regressions of Future Abnormal Returns on the Ranks of DuPont Components

$$R_{t+1} = \rho_0 + \rho_1 \Delta RNOA_t + \rho_2 \Delta PM_t + \rho_3 \Delta ATO_t + \rho_4 RSST\ Controls + \rho_5 RNOA_t + \rho_6 PM_t + \rho_7 ATO_t + \rho_8 Fama\ French\ Risk\ Factors + \varepsilon_{t+1}$$

Independent Variables	**Model 1**	**Model 2**	**Model 3**
Intercept	−0.061	−0.014	−0.115
	−2.15	−0.49	−2.83
$\Delta RNOA_t$	0.001	−0.001	−0.000
	0.11	−0.12	−0.08
ΔPM_t	0.006	0.016	0.016
	0.33	1.23	0.73
ΔATO_t	0.078	0.054	0.052
	5.12	3.11	2.52
ΔWC_t		−0.513	−0.557
		−4.61	−5.58
ΔNCO_t		−0.162	−0.193
		−3.96	−5.26
ΔFIN_t		−0.041	−0.108
		−1.14	−2.29
$RNOA_t$			0.070
			1.41
PM_t			0.110
			1.04
ATO_t			0.000
			0.29
FF Risk Factors	Included	Included	Included
Adjusted R^2	1.6%	3.0%	3.8%

The sample consists of 38,716 firm-year observations from 1984–2002. Regression results are based on 19 annual regressions using the Fama-MacBeth approach of averaging coefficients and calculating the t-statistics as follows: $\hat{\rho} \Big/ \frac{StdDev\rho}{\sqrt{n}}$ where $\hat{\rho}$ is the simple average of the annual coefficients, n is the number of annual regressions and $StdDev\rho$ is the standard deviation of the annual coefficients.

R_{t+1} (*Future Abnormal Returns*) = compound buy-and-hold returns inclusive of all dividends and other distributions less the value-weighted market index beginning four months after the end of the fiscal year and continuing for one year; variables are ranked annually and assigned to deciles; the continuous value of the variables is replaced by decile-rank in the regressions and divided by 9;

NOA (*Net Operating Assets*) = Operating Assets − Operating Liabilities; Operating Assets is calculated as total assets (item #6) less cash and short-term investments (item #1) and (item #32); Operating Liabilities is calculated as total assets less total debt (items #9 and #34), less book value of total common and preferred equity (items #60 and #130), less minority interest (item #38);

PM_t = Operating Income$_t$ (item #178)/Total Sales$_t$ (item #12);

ATO_t = Sales$_t$/Average NOA_t $((NOA_t + NOA_{t-1})/2)$;

$RNOA_t = PM_t \times ATO_t$;

$\Delta PM_t = PM_t - PM_{t-1}$;

$\Delta ATO_t = ATO_t - ATO_{t-1}$;

$\Delta RNOA_t = RNOA_t - RNOA_{t-1}$;

(*continued on next page*)

TABLE 7 (continued)

ΔWC_t = $WC_t - WC_{t-1}$; *WC* is calculated as Current Operating Assets (COA) − Current Operating Liabilities (COL), and COA = Current Assets (Compustat Item #4) − Cash and Short-Term Investments (STI) (Compustat Item #1), and COL = Current Liabilities (Compustat Item #5) − Debt in Current Liabilities (Compustat Item #34).

ΔNCO_t = change in net noncurrent operating assets is defined as $NCO_t - NCO_{t-1}$. NCO is calculated as Non-Current Operating Assets (NCOA) − Non-Current Operating Liabilities (NCOL), and NCOA = Total Assets (Compustat item #6) − Current Assets (Compustat Item #4) − Investments and Advances (Compustat Item #32), and NCOL = Total Liabilities (Compustat Item #181) − Current Liabilities (Compustat Item #5) − Long-Term Debt (Compustat Item #9).

ΔFIN_t = change in net financial assets is defined as $FIN_t - FIN_{t-1}$, and FIN = Financial Assets (FINA) − Financial Liabilities (FINL). FINA = Short-Term Investments (STI) (Compustat Item #193) + Long-Term Investments (LTI) (Compustat Item #32), and FINL = Long-Term Debt (Compustat Item #9) + Debt in Current Liabilities (Compustat Item #34) + Preferred Stock (Compustat Item #130).

FF Risk Factors are comprised of book-to-market ratio, log (MVE), and Beta. *Book-to-Market Ratio* = Book value of equity (Compustat item #60)/Market Value of Equity (shares × price − item #25 × item #199). *MVE* = Log (Market Value of Equity). *BETA* (β) for firm *i* for fiscal year *t* is estimated by a market model regression. The regression is run using weekly returns for a period of two years ending at the end of the fiscal year from which I obtain the data to compute each of the financial ratios.

about 7.8 percent after controlling for $\Delta RNOA$ and the three Fama and French (1993) risk factors. The return drops to about 5.2 percent but remains statistically significant, even after controlling for the RSST accrual measures and the levels of the DuPont components. Thus, market participants do not appreciate the persistent nature of changes to asset efficiency as measured by *ATO*.

Analyst Forecast Revision Tests

Table 8 presents the first set of analyst tests that are used to ascertain whether analyst forecast revisions are associated with the DuPont components. In the first regression, the forecast surprise is positive and significant, consistent with analysts revising their forecasts of future earnings in the same direction as the earnings surprise. Model 2 adds the DuPont components. Once again, ΔATO is the only DuPont component that is positive and significant. Model 3 shows that the significance remains even after including the change of *RNOA*.[33] Recall that changes in *ATO* positively predict future changes in profitability. Thus, the association between analyst forecast revisions and changes in the DuPont components is consistent with their predictive ability of future profitability and indicates that analysts are revising their forecasts in a manner consistent with the predictive properties of ΔATO.

Table 9 presents the analysis of explaining future forecast errors from Equation (8). Model 1 presents the levels of the components. Both are insignificant, which is again consistent with the lack of power on predicting future profitability. Model 2 presents the changes specification. Surprisingly, both ΔPM and ΔATO are statistically significant. This is surprising because ΔPM (1) has no predictive power with respect to future changes in profitability and (2) generally has similar time-series properties to *RNOA* (Nissim and Penman 2001). Accordingly, one would expect analysts to properly impound the information in this variable.

---

[33] In unreported portfolio tests, ΔPM also appeared to have some predictive power when the earnings surprise control was not included.

TABLE 8
Time-Series Means and t-Statistics for Coefficients from Annual Cross-Sectional Regressions of Analysts Forecast Revisions on the DuPont Components

$$Anal\_REV_t = \rho_0 + \rho_1 SUR_t + \rho_2 \Delta PM_t + \rho_3 \Delta ATO_t + \rho_4 \Delta RNOA_t + \varepsilon_t$$

	Intercept	SUR_t	ΔPM_t	ΔATO_t	$\Delta RNOA_t$	**Adj. R^2**
Model 1	−0.002	0.062				1.09%
	(−5.40)	(3.06)				(2.66)
Model 2	−0.002	0.066	0.001	0.001		1.4%
	(−4.44)	(3.26)	(1.09)	(3.63)		(2.88)
Model 3	−0.002	0.066	0.001	0.001	0.000	1.32%
	(−4.40)	(3.26)	(0.94)	(2.44)	(0.01)	(3.02)

The sample consists of 38,716 firm-year observations from 1984–2002. Regression results are based on 19 annual regressions using the Fama-MacBeth approach of averaging coefficients and calculating the t-statistics as follows: $\hat{\rho} \Big/ \frac{StdDev\rho}{\sqrt{n}}$ where $\hat{\rho}$ is the simple average of the annual coefficients, n is the number of annual regressions and $StdDev\rho$ is the standard deviation of the annual coefficients.

$Anal\_REV_t$ (*Analyst Revision*) = one-year-ahead earnings forecasts measured as the revision to the consensus analyst forecast of year $t+1$ earnings made just after year t earnings are announced. Specifically, it is the first median I/B/E/S consensus one-year-ahead forecast of year $t+1$ earnings, minus the last median consensus of year $t+1$ earnings made directly before the announcement of year t earnings, all scaled by share price at the end of fiscal year $t-1$;

SUR_t (*Annual Earnings Surprise*) = annual I/B/E/S earnings − the most recent median forecast of annual earnings for year t;

NOA (*Net Operating Assets*) = Operating Assets − Operating Liabilities; Operating Assets is calculated as total assets (item #6) less cash and short-term investments (item #1) and (item #32); Operating liabilities is calculated as total assets less total debt (items #9 and #34), less book value of total common and preferred equity (items #60 and #130), less minority interest (item #38);

ΔNOA (*Growth in Net Operating Assets*) = $(NOA_t - NOA_{t-1})/NOA_{t-1}$;

PM_t = Operating Income$_t$ (item #178)/Total Sales$_t$ (item #12);

ATO_t = Sales$_t$/Average NOA_t $((NOA_t + NOA_{t-1})/2)$;

$RNOA_t = PM_t \times ATO_t$;

$\Delta PM_t = PM_t - PM_{t-1}$;

$\Delta ATO_t = ATO_t - ATO_{t-1}$; and

$\Delta RNOA_t = RNOA_t - RNOA_{t-1}$.

But the analysis suggests that analysts do not appear to employ all the information in ΔATO when making forecasts. The positive sign of the coefficient is also important. Recall that ΔATO positively predicts future changes in profitability after controlling for the $\Delta RNOA$. The positive coefficient on this variable indicates that, on average, analysts' forecast errors seem to be predicted by ΔATO, consistent with the positive prediction of future profitability and is entirely consistent with the future abnormal return tests. Specifically, it gives support to the idea that analysts do not fully appreciate the information in this variable, and that (similar to Abarbanell and Bernard [1992]) market participants may be relying on analysts for information. Overall, the two tests seem to confirm that there is information contained in ΔATO that is not completely understood by market participants, and both point to the lack of appreciation of the predictive abilities of ΔATO.

TABLE 9

Time-Series Means and t-Statistics for Coefficients from Annual Cross-Sectional Regressions of Future Forecast Errors on the DuPont Components

$$FE_{t+1} = \rho_1 + \rho_2 \Delta RNOA_t + \rho_3 PM_t + \rho_4 ATO_t + \rho_5 \Delta PM_t + \rho_6 \Delta ATO_t + \varepsilon_{t+1}$$

	Intercept	$\Delta RNOA_t$	PM_t	ATO_t	ΔPM_t	ΔATO_t	Adj. R^2
Model 1	0.013	−0.001	−0.013	−0.000			1.74%
	(8.44)	(−1.35)	(−1.62)	(−0.09)			(1.46)
Model 2	0.015	0.001			0.002	0.002	2.78%
	(11.27)	(1.09)			(3.29)	(2.56)	(7.45)
Model 3	0.013	0.000	−0.013	−0.000	0.002	0.002	1.79%
	(8.57)	(0.48)	(−1.83)	(−0.30)	(3.27)	(2.54)	(8.32)

The sample consists of 38,716 firm-year observations from 1984–2002. Regression results are based on 19 annual regressions using the Fama-MacBeth approach of averaging coefficients and calculating the t-statistics as follows: $\bar{\rho} \Big/ \frac{StdDev\rho}{\sqrt{n}}$ where $\bar{\rho}$ is the simple average of the annual coefficients, n is the number of annual regressions, and $StdDev\rho$ is the standard deviation of the annual coefficients.

FE_{t+1} (*Forecast error*) = realized earnings for year $t+1$ less the median forecast earnings from the months prior to the announcement of $t+1$ earnings, scaled by the stock price in the month of the earnings announcement for year t;

NOA (*Net Operating Assets*) = Operating Assets − Operating Liabilities; Operating Assets is calculated as total assets (item #6) less cash and short-term investments (item #1) and (item #32); Operating liabilities is calculated as total assets less total debt (items #9 and #34), less book value of total common and preferred equity (items #60 and #130), less minority interest (item #38);

PM_t = Operating Income$_t$ (item #178)/Total Sales$_t$ (item #12);

ATO_t = Sales$_t$/Average NOA_t $((NOA_t + NOA_{t-1})/2)$;

$RNOA_t = PM_t \times ATO_t$;

$\Delta PM_t = PM_t - PM_{t-1}$;

$\Delta ATO_t = ATO_t - ATO_{t-1}$; and

$\Delta RNOA_t = RNOA_t - RNOA_{t-1}$.

V. CONCLUSION

In this paper, the usefulness of DuPont analysis, a common form of financial statement analysis that decomposes return on net operating assets (*RNOA*) into two multiplicative components: profit margin and asset turnover, is comprehensively tested. This decomposition breaks down earnings into two meaningful ratios that are interpretable. Consistent with prior literature and the economic intuition that an increase in asset turnover represents an improvement in asset utilization that will generate future results, I find that changes in asset turnover have positive explanatory power for future changes in *RNOA* incremental to current changes in *RNOA*. This paper contributes to this growing literature by confirming that the information contained in changes in asset turnover is incremental to other documented predictors of profitability such as accruals and the fundamental signals used by experts in Lev and Thiagarajan (1993).

However, the primary focus of the paper is how market participants, such as equity analysts and stock market investors, use DuPont components in assessing the prospects of the firm. The paper examines the stock market's association with the information in the DuPont components variables through long- and short-window tests and finds that the market recognizes the future *RNOA* implications of these components. As would be expected, long-window stock returns are positively associated with changes in asset turnover. Additionally, short-window returns show that the market responds favorably to changes in asset turnover. However, future return tests confirm that this favorable response is incomplete. Analyst forecast revision tests mirror the short-window return results. I find that there is positive revision to future earning forecasts associated with changes in asset turnover that is incremental to the earnings surprise as well as changes in total profitability. Thus, sell-side analysts appear to properly use the information in ΔATO in assessing the future prospects of the firm at least from a directional perspective. However, I find that although there is an association with this ratio and revisions of future earnings forecasts, there appears to be information in the DuPont components that goes unused by analysts as evidenced by the predictability of future forecast errors. The consistency of the future analyst forecast error tests with the future return tests is reassuring. In the spirit of Abarbanell and Bernard (1992), it points to the possibility that investors may be misled partially because analysts do not impound the differential time-series properties of these two components.

Overall, the study finds that the DuPont analysis is a useful tool of financial statement analysis. Unlike other analyses, the DuPont decomposition of *RNOA* has the nice feature of being derived from a theoretical and parsimonious framework of valuation and relates to the operational aspects of the firm. In particular, the analysis points to the value of focusing on changes in asset turnover.

REFERENCES

Abarbanell, J. 1991. Do analysts' earnings forecasts incorporate information in prior stock price changes? *Journal of Accounting and Economics* 14: 147–165.

———, and V. Bernard. 1992. Test of analysts' overreaction/underreaction to earnings information as an explanation for anomalous stock price behavior. *Journal of Finance* 47 (July): 1181–1207.

———, and B. Bushee. 1997. Fundamental analysis, future earnings, and stock prices. *Journal of Accounting Research* (Spring): 1–24.

———, and ———. 1998. Abnormal returns to a fundamental analysis strategy. *The Accounting Review* (January): 19–45.

Alford, A. W., J. Jones, and M. Zmijewski. 1994. Extensions and violations of the statutory SEC Form 10-K filing requirements. *Journal of Accounting and Economics* 17: 229–256.

Amir, E., and B. Lev. 1996. Value-relevance of non-financial information: The wireless communication industry. *Journal of Accounting and Economics* 22: 3–30.

Ball, R., and P. Brown. 1968. An empirical evaluation of accounting numbers. *Journal of Accounting Research* (Autumn): 159–177.

Barth, M., W. Beaver, J. Hand, and W. Landsman. 1999. Accruals, cash flows and equity values. *Review of Accounting Studies* (December): 205–229.

———, and R. Hutton. 2004. Analyst earnings forecast revisions and the pricing of accruals. *Review of Accounting Studies* (March): 59–96.

Beaver, W. 1968. The information content of annual earnings announcements. *Journal of Accounting Research* (Supplement): 67–92.

———, P. Kettler, and M. Scholes. 1970. The association between market-determined and accounting-determined risk measures. *The Accounting Review* (October): 654–682.

———, R. Lambert, and D. Morse. 1980. The information content of security prices: A second look. *Journal of Accounting and Economics* 9: 139–158.

———. 1998. *Financial Reporting: An Accounting Revolution.* 3rd edition. Upper Saddle River, NJ: Prentice Hall.

Bernard, V. 1987. Cross-sectional dependence and problems in inference in market-based accounting research. *Journal of Accounting Research* (Spring): 1–48.

———, and J. Thomas. 1989. Post-earnings announcement drift: Delayed price response or risk premium? *Journal of Accounting Research* (Supplement): 1–36.

Bradshaw, M., S. Richardson, and R. Sloan. 2001. Do analysts and auditors use information in accruals? *Journal of Accounting Research* (Summer): 45–74.

Brown, S., K. Lo, and T. Lys. 1999. Use of R^2 in accounting research: Measuring change in value relevance over the last four decades. *Journal of Accounting and Economics* 28: 83–115.

Burgstahler, D., and I. Dichev. 1997. Earnings, adaptation and equity values. *The Accounting Review* (April): 187–215.

Chen, S., M. DeFond, and C. Park. 2002. Voluntary disclosure of balance sheet information in quarterly earnings announcements. *Journal of Accounting and Economics* 33: 229–251.

Collins, D., S. P. Kothari, J. Shanken, and R. Sloan. 1994. Lack of timeliness versus noise as explanations for low contemporaneous return-earnings association. *Journal of Accounting and Economics* 18: 289–324.

———, M. Pincus, and H. Xie. 1999. Equity valuation and negative earnings: The role of book value of equity. *The Accounting Review* (January): 29–61.

Doyle, J., R. Lundholm, and M. Soliman. 2003. The predictive value of expenses excluded from pro forma earnings. *Review of Accounting Studies* 8: 145–174.

Easton, P., and T. Harris. 1991. Earnings as an explanatory variable for returns. *Journal of Accounting Research* (Spring): 19–36.

Elgers, P., M. Lo, and R. Pfeiffer. 2001. Delayed security price adjustments to financial analysts' forecasts of annual earnings. *The Accounting Review* 76 (4): 613–632.

———, ———, and ———. 2003. Analysts vs. investors weightings of accruals in forecasting annual earnings. *Journal of Accounting and Public Policy* 22: 255–280.

Fairfield, P., R. Sweeney, and T. Yohn. 1996. Accounting classification and the predictive content of earnings. *The Accounting Review* (July): 337–355.

———, and T. Yohn. 2001. Using asset turnover and profit margin to forecast changes in profitability. *Review of Accounting Studies* (December): 371–385.

———, J. Whisenant, and T. Yohn. 2003a. The differential persistence of accruals and cash flows for future operating income versus future profitability. *The Review of Accounting Studies* 8: 221–243.

———, ———, and ———. 2003b. Accrued earnings and growth: Implications for earnings persistence and market mispricing. *The Accounting Review* (January): 353–371.

Fama, E., and J. Macbeth. 1973. Risk, return and equilibrium—Empirical tests. *Journal of Political Economy* (May/June): 607–641.

———. 1990. Stock returns, expected returns and real activity. *Journal of Finance* (September): 1089–1108.

———, and K. French. 1993. Common risk factors in the returns of stocks and bonds. *Journal of Financial Economics* 33: 3–56.

Feltham, G., and J. Ohlson. 1995. Valuation and clean surplus accounting for operating and financial activities. *Contemporary Accounting Research* (Spring): 689–731.

Francis, J., and K. Schipper. 1999. Have financial statements lost their relevance? *Journal of Accounting Research* (Autumn): 319–352.

Ge, W., and M. Soliman. 2007. Achieving competitive advantage through product differentiation and cost efficiency: A new approach with new insights. Working paper, University of Washington.

Hayn, C. 1995. The information content of losses. *Journal of Accounting and Economics* 20 (2): 125–154.

Holthausen, R., and D. Larcker. 1992. The prediction of stock returns using financial statement analysis. *Journal of Accounting and Economics* 15: 373–411.

Iman, R., and W. Conover. 1979. The use of the rank transformation in regression. *Technometrics* (November): 499–509.

Klein, A. 1990. A direct test of the cognitive bias theory of share price reversals. *Journal of Accounting and Economics* 13: 155–166.

Kormendi, R., and R. Lipe. 1987. Earnings innovations, earnings persistence, and stock returns. *Journal of Business* 60: 323–345.

Kothari, S. P. 1992. Price-earnings regressions in the presence of prices leading earnings. *Journal of Accounting and Economics* 15: 173–202.

———, and R. Sloan. 1992. Information in prices about future earnings: Implications for earnings response coefficients. *Journal of Accounting and Economics* 15: 143–171.

———. 2001. Capital markets research in accounting. *Journal of Accounting and Economics* 31: 105–231.

Lee, C. 1999. Accounting-based valuation: Impact on business practices and research. *Accounting Horizons* 13: 413–425.

Lev, B. 1989. On the usefulness of earnings: Lessons and directions from two decades of empirical research. *Journal of Accounting Research* (Supplement): 153–192.

———, and S. Thiagarajan. 1993. Fundamental information analysis. *Journal of Accounting Research* (Autumn): 190–215.

Lipe, R. 1986. The information contained in the components of earnings. *Journal of Accounting Research* (Supplement): 37–64.

Lys, T., and S. Sohn. 1990. The association between revisions of financial analysts' earnings forecasts and security price changes. *Journal of Accounting and Economics* 13: 341–363.

McNichols, M., and P. O'Brien. 1997. Self-selection and analyst coverage. *Journal of Accounting Research* (Supplement): 167–199.

Mendenhall, R. 1991. Evidence on the possible underweighting of earnings-related information. *Journal of Accounting Research* 29 (10): 170–179.

Modigliani, F., and M. Miller. 1958. The cost of capital, corporation finance and the theory of investment. *American Economic Review* (June): 261–297.

Newey, W., and K. West. 1987. A simple, positive, semi-definite, heteroscedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55: 703–708.

Nissim, D., and S. Penman. 2001. Ratio analysis and equity valuation: From research to practice. *Review of Accounting Studies* (March): 109–154.

Ohlson, J. 1995. Earnings, book values, and dividends in security valuation. *Contemporary Accounting Research* (Spring): 661–687.

———. 1999. On transitory earnings. *Review of Accounting Studies* (December): 145–162.

Ou, J., and S. Penman. 1989. Financial statement analysis and the prediction of stock returns. *Journal of Accounting and Economics* 11: 295–329.

Penman, S. 1992. Return to fundamentals. *Journal of Accounting, Auditing and Finance* (Autumn): 465–482.

———. 2001. *Financial Statement Analysis and Security Valuation.* 1st edition. New York, NY: McGraw Hill.

———, and X. Zhang. 2003. Modeling sustainable earnings and P/E ratios using financial statement information. Working paper, Columbia University.

Richardson, S., R. Sloan, M. Soliman, and I. Tuna. 2005. Accrual reliability, earnings persistence and stock prices. *Journal of Accounting and Economics* 39: 437–485.

———, ———, ———, and ———. 2006. The implications of firm growth and accounting distortions for accruals and profitability. *The Accounting Review* 81: 713–743.

Romer, P. 1986. Increasing returns and long-run growth. *Journal of Political Economy* 94 (51): 1002–1037.

Shumway, T. 1997. The delisting bias in CRSP data. *Journal of Finance* 52: 327–340.

———, and V. Warther. 1999. The delisting bias in CRSP's Nasdaq data and its implications for the size effect. *Journal of Finance* 54: 2361–2379.

Sloan, R. 1996. Do stock prices fully reflect information in accruals and cash flows about future earnings? *The Accounting Review* (July): 289–315.

Soliman, M. 2004. Using industry-adjusted DuPont analysis to predict future profitability and returns. Ph.D. dissertation, University of Michigan.

Stober, T. 1992. Summary financial statement measures and analysts' forecasts of earnings. *Journal of Accounting and Economics* 15: 347–372.

Verbeek, M. 2000. *A Guide to Modern Econometrics.* 1st edition. New York, NY: John Wiley & Sons.

Vuong, Q. 1989. Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica* (March): 307–333.

Walther, B. 1997. Investor sophistication and market earnings expectations. *Journal of Accounting Research* 35 (2): 157–179.

Warfield, T. and J. Wild. 1992. Accounting recognition and the relevance of earnings as an explanatory variable for returns. *Accounting Review* 67 (4): 821–842.

THE ACCOUNTING REVIEW
Vol. 83, No. 3
2008
pp. 855–863

BOOK REVIEWS

Stephen A. Zeff, Editor

Editor's note: Two copies of books for review should be sent to the Book Review Editor: Stephen A. Zeff, Rice University, Jesse H. Jones Graduate School of Management, 6100 Main St., Houston, TX 77005. The policy of *The Accounting Review* is to publish only those reviews solicited by the Book Review Editor. Unsolicited reviews will not be accepted.

RICK ANTLE, FRØYSTEIN GJESDAL, and PIERRE JINGHONG LIANG (editors), *Essays in Accounting Theory in Honour of Joel S. Demski* (New York, NY: Springer, 2007, ISBN 13: 978-0387-30397-0, pp. xix, 326).

"Since the 1960s, Joel Demski has been a potent force in the accounting scholarship and education. Always a rebel, Joel sees any set of received doctrines as more of a target (in the military sense) than as being set in stone" (p. xvii). These sentences form the promising beginning of the Preface of a book that would be more typically seen in Continental Europe than in the U.S. It is a book that may best be characterized as a *Festschrift*, that is, a collection of articles with the following common attributes: the editors are closely related to the honored scholar, and here they are former doctoral students of Joel Demski; the contributing authors have a relationship to the scholar in one way or the other; and the book contains a list of his publications. One article surveys Joel's contributions to accounting research and places his work in perspective, while the other articles reflect Joel's research interests and often explicitly motivate their analysis by referring to his work. Most of the articles use analytical methods that Joel was instrumental in developing and promoting in accounting research. The only characteristic of the book that fails to be consistent with a *Festschrift* is the fact that it was published in early 2007, a year with no obvious connection to a major birthday or another "*Fest*" of the honored. The Preface, however, is dated September 2005 and mentions that the book was a joint effort with a conference held in October 2005, which coincided with Joel's 65th birthday. So, the book appears to have been published somewhat late, although this certainly does not diminish the value of its contents, which are quite topical.

The book contains 14 previously unpublished articles that were written by 25 highly regarded authors. The first article by Jerry Feltham is an overview of Joel's leading role in promoting accounting scholarship. It is a personal reflection of the origins and development of Joel's work, much of which was the joint work of Joel and Jerry. This article is useful for understanding the background of the development of what was then a new theory in accounting research that started in the Sixties and which soon became generally accepted.

The rest of the book is divided into four parts. As the book is a diverse collection of articles, these parts can provide only a loose structure for guiding the reader through the articles. Within each part, the articles are apparently ordered alphabetically by author. I provide a brief description of these parts.

The first part, "General Theory," contains five articles. John Christensen and Hans Frimor find that incorporating fair value accounting information in a rational expectations equilibrium does not necessarily increase market informativeness in the light of other sources of information. Ron Dye and Mark Finn study voluntary disclosure and show another reason why the usual unraveling breaks down and nondisclosure can prevail in equilibrium. The reason is the randomness of the firm's information quality, which they measure by sample size. John Fellingham and Doug Schroeder explore a setting in which synergies in activities are modeled using the structure of quantum probabilities. They show that the aggregate measurement of such bundled activities is more efficient than individual measurement. Frøystein Gjesdal analyzes optimal contracting and the value of information in a moral hazard setting with hidden information, where the principal can observe the action but not the conditions on which the agent's action is based. Finally, Thomas Hemmer attacks the perception of many researchers that agency theory predicts a negative relation between risk and the sensitivity of pay-for-performance. He shows that, except for special cases, there is no such directional prediction, thus providing a theoretical explanation for the weak empirical support of a negative relation.

The second part, "Applied Theory," includes four articles. Rick Antle, Peter Bogetoft, and Andy Stark study a capital rationing and organizational slack setting in which a privately informed manager can influence two decisions, whether or not to invest and whether to delay investment. They show the complex effects of this additional flexibility on the optimal cost targets. Anil Arya, Jon Glover, and Brian Mittendorf examine why many firms cap bonus payments to managers even though the caps are usually considered to be non-optimal. They find that caps complement linear contracts in aligning impatient managers' incentives with respect to investment decisions. In another article, Anil Arya, Jon Glover, and Suresh Radhakrishnan take a closer look at the controllability principle. They provide examples of agency settings in which a signal is informative but not valuable in contracting and in which a signal is uninformative but valuable. Steve Huddart, Jack Hughes, and Carolyn Levine study a capital market equilibrium where insiders trade on common private information. They show that the public disclosure of insider trading may help them better coordinate their actions at the expense of liquidity traders and lead to less price discovery: private disclosure can therefore be preferable.

The third part, "Connections to Practice," contains two articles. Joseph Gerakos, Chris Ittner, and Dave Larcker study the use of performance-vested stock options rather than traditional stock options awarded to executives in U.S. firms. They describe determinants for their use and the choice of their characteristics. William Rogerson provides a fictional case study illustrating incentives for the distortion of cost allocations performed by defense contractors as between government and private business. These two articles stand in contrast to the otherwise highly homogeneous articles in terms of fundamental theory and method.

The fourth and last part, "Commentary and Perspective," includes two articles that lead the reader back to more general views of accounting and accounting research. Gary Sundem pleads for a re-examination of the conceptual framework in the light of insights provided by the information content view of accounting. Finally, Shyam Sunder reflects on the economizing principle borrowed from the natural sciences and illustrates its application to various accounting issues. He closes by recounting personal encounters with Joel Demski, which provide a neat end to the book that began with Feltham's personal views.

The back of the book mentions established and aspiring accounting scholars as the target (not in a military sense) audience, and I agree with that. It might be useful to add that analytically inclined scholars would certainly benefit more than others. Most of the articles are "hard core" analytical articles that make no concessions to the readers' background knowledge of analytical methods and the related research. Some articles, though, are different either in the research method used or are broader in perspective and are accessible also by a more generally interested audience. I am not sure if this heterogeneity was originally intended by the editors, but it is probably not a bad thing to have a broader scope for a book such as this.

Using the common theme in the book—information asymmetry and incentives—one may wonder whether the fact that the articles are published in this book rather than in journals is a signal about the expected usefulness of reading them. They are certainly of high quality, with the reputation of the highly respected authors standing behind them. Most articles are probably of interest only to a smaller or specialized community. However, the book makes up for that because it includes a diverse set of topics, thus covering a broad spectrum of themes. I am sure everyone in the target audience will find several of the articles beneficial and useful to digest. For example, I found the articles in the first part of the book particularly interesting. The remaining articles give a good overview of research issues that are examined by the use of analytical research.

The book is therefore highly recommended to accounting researchers and doctoral students who are interested in accounting scholarship—the very objective Joel Demski has always strived to foster.

ALFRED WAGENHOFER
Professor of Accounting
Karl-Franzens-Universität Graz

ALNOOR BHIMANI (editor), *Contemporary Issues in Management Accounting* (Oxford, U.K.: Oxford University Press, 2006, ISBN 978-0-19-928335-4, pp. xvi, 447).

This book contains 18 chapters, each of which is a paper on some aspect of management accounting. The goal of the book was to capture "key facets of current thoughts, concerns, and issues in management accounting" (p. vii). The authors are prominent academics from around the world. The book was dedicated to Professor Michael Bromwich, to recognize his many contributions to management accounting thought and practice on the occasion of Michael's retirement from the faculty of the London School of Economics.

This is not a book that one would pick up and read from front to back. The book has no theme except, as the title indicates, contemporary issues in management accounting. The topics are quite disparate. Most of the chapters are reflective essays on the state of knowledge and research in a specific area of management accounting. No attempt is made to organize the papers by topics; the papers are ordered alphabetically by last name of first author.

Maybe the best way to describe what it is in a book without a single theme is by providing a short synopsis of each chapter. I will do so with an attempt to group the chapters by broad topic area.

Quite a few of the chapters deal with specific aspects of management control systems (MCSs). In Chapter 11, Kim Langfield-Smith provides a review of much of the literature focusing on understanding the links between strategy and MCSs. She also describes how research approaches in this area have evolved over the past almost 50 years.

Two chapters explicitly focus on a major MCS element: performance measurement. In Chapter 1, Thomas Ahrens and Christopher Chapman discuss links between performance metrics and activities within the firm and present some thoughts on how academic research can contribute to the positive aspects of those links, while minimizing the negative. In Chapter 5, Robert Chenhall reviews recent developments in performance measures (e.g., economic value, nonfinancial, and integrated performance). He then explains how some frequently studied situational contingencies (external environment and strategy, technology, structure, and size) might influence the effectiveness of each of these new measures.

In Chapter 2, Stanley Baiman provides a nontechnical explanation of the contract theory model, which, he asserts, has become "the dominant analytical research paradigm in management accounting" in the last two decades (p. 20). He then describes some examples of how the model can be applied to some common management accounting issue areas, including variance investigation and the use of stock prices and subjectivity for performance evaluation and compensation purposes.

David Otley's focus in Chapter 13 is on other MCS elements: budgetary control and responsibility accounting. He reviews the many roles that budgets play in organizations and the problems they often cause. Then he discusses two different approaches for minimizing the problems: improving the budgetary processes or adapting the more radical approach of going "Beyond Budgeting."

Three of the chapters discuss specific aspects of cost accounting. In Chapter 6, Robin Cooper and Regine Slagmulder describe "integrated cost management programs" wherein the output of one cost management technique becomes the input for another technique. They describe the findings of their field research done in Japanese firms that discovered five "internal" (i.e., within a single entity) cost management techniques: (1) target costing, (2) product-specific Kaizen costing, (3) general Kaizen costing, (4) functional group management, and (5) standard costing. They also discuss five "external" (i.e., linking the firm with buyers and suppliers) techniques: (1) functional-price-quality trade-offs, (2) inter-organizational cost investigations, and (3) concurrent cost management. The chapter provides examples and discussions of when these approaches should be used.

In Chapter 10, Eva Labro discusses the findings of analytical and empirical research on the design of cost systems. She reviews cost system objectives, and describes both the approximations that traditional and activity-based cost systems make and the problems that these approximations can cause. Then she describes how to choose a costing system and concludes with some suggestions for advancing costing system research and practice.

In Chapter 16, the late John Shank provides a 30-year history of a field that, he believes, should have been transformed from a focus on cost accounting to management accounting to strategic accounting. But, he laments, the promise of strategic accounting or strategic cost management never came to fruition for a number of reasons, the most important of which might be the inability, or unwillingness, of practicing accountants to think strategically.

While many of the chapters provide a focused literature review in a specific area of management accounting, two chapters, in addition to the Shank chapter, provide broader historical analyses. In Chapter 15, Robert Scapens traces developments and trends in the management accounting field, particularly in the U.K., over the 35 years of his career. He describes and celebrates the trend toward greater pluralism—theoretical and methodological diversity—in management accounting, but he laments the lack of a significant impact on practice. In Chapter 8, Keith Hoskin, Richard Macve, and John Stone provide a re-analysis of the history of the concept of strategy, in both business and military organizations, going back many hundreds of years. They explain that accounting is a critical enabling technology. It is used to evaluate ends and means and to hold individuals accountable for performance, thereby objectifying and rationally controlling the future.

Two of the chapters describe and analyze management accounting practices in a specific industry. In Chapter 9, Liisa Kurunmäki and Peter Miller discuss the "modernizing government" agenda in the U.K. They found many similarities with private practice, because many of the U.K. government's efforts involved attempts to utilize formalized systems of governance and performance assessment. But they also found some major differences and impediments, particularly because of the government agencies' needs for cooperative working patterns. In Chapter

18, Mark Young, Wim Van der Stede, and James Gong explain how the domain of management accounting can be expanded into the creative industries. Their detailed example is of the U.S. motion picture industry.

The topics of the remaining chapters are relatively independent except in a broad sense. In Chapter 7, Lawrence Gordon, Martin Loeb, and Chih-Yang Tseng discuss developments in the practice of capital budgeting. They conclude that the focus in the literature has been on improving the models used to rank and select specific capital projects (e.g., real options models). They call for management accounting researchers to devote greater attention to the relatively understudied control side of the capital budgeting process, which involves both getting good information into the models and holding managers accountable for executing their plans.

In Chapter 3, Jane Baxter and Wai-Fong Chua conclude that there is more to management accounting than just economically rational planning and control. They provide seven ways of viewing (or framing) the effects of management accounting in practice (e.g., naturalistic frame, structurationist frame, institutional frame). They explain how these alternate frames provide a richer understanding of how management accounting does and should work in practice.

In Chapter 4, Alnoor Bhimani, the editor, discusses some of the impacts of advancing information technologies, what he calls "digitization," on management accounting practice. Some of these impacts are direct, because the costs of collecting, manipulating, and reporting information have declined dramatically. Others are indirect, as the new technologies have led to the creation of new sources of value creation and new business models.

In Chapter 12, Jan Mouritsen and Allan Hansen discuss management challenges at the operational level, particularly in modern lean manufacturing and extended supply chain environments together with contributions that management accounting can make. They conclude that, while operations managers rely heavily on nonfinancial performance indicators, accounting indicators also play important roles. But accounting must continue to adapt to the changing operational settings.

In Chapter 14, Hanno Roberts takes a resource-based theory perspective of the firm in which the firm is primarily a mechanism for the production of knowledge. He then discusses the role of management accounting in helping to coordinate and integrate the accumulated and generated knowledge.

In Chapter 17, Kazbi Soonawalla describes the importance and developing practice of environmental management accounting.

Stepping back from all the specifics, anyone who picks up this book will be struck by the great diversity of topics that are included under the broad rubric "management accounting." The chapters in this book do not discuss all the possibilities, but they do demonstrate the breadth of the field.

All of the chapters in this book are well written. Many of the chapters provide a concise introduction to a part of the management accounting field, a snapshot of the current state of knowledge, and both predictions of and suggestions for the future.

Which chapters would I particularly recommend? Readers' preferences will vary significantly depending on their topical interests and the purposes for which the chapters are intended. For someone like me with a primary interest in MCS, the reviews of the MCS literature were less interesting because I already understood that literature well. But I may well use one or more of these chapters as a required reading in an MCS-related doctoral seminar or even an M.B.A. or advanced undergraduate class.

Quite a few chapters gave me ideas for possible future research. For example, I made a number of notes while reading the Ahrens and Chapman chapter (#1). And I was intrigued by one section, particularly, of the Baiman chapter (#2). Baiman uses a contract theory model to argue that compensating managers based on stock price performance is not optimal because the aggregation of information signals for price estimation purposes is quite different from that for performance measurement purposes. I had never thought about this important issue in this way. I am still contemplating whether I need to change my lectures on this topic or to write a paper explaining why this new Baiman insight is, in the grand scheme of things, not particularly important.

My reading of many of the other chapters educated me in a more general sense. For example, I found some of the historical and literature review chapters in the non-MCS topic areas very informative. I will definitely use some of the material in the cost accounting and capital budgeting review chapters to inform my teaching. And, just to provide another example, the Baxter and Chua chapter (#3) was an eye-opener for me. It helped give me a basic understanding of the importance of many relevant bodies of research, or "frames," with which I (and most other U.S.-based accounting researchers) have little familiarity.

The material in this book is as rich as it is diverse. The book will not appeal to anyone in its entirety, but parts of it will appeal to almost everyone interested in management accounting.

KENNETH A. MERCHANT
Professor of Accounting
University of Southern California

JAYNE M. GODFREY and KERYN CHALMERS (editors), *Globalisation of Accounting Standards*, Monash Studies in Global Movements (Cheltenham, U.K.: Edward Elgar, 2007, ISBN 978-1-84542-852-5, pp. xvi, 309).

It is hardly surprising that a book entitled *Globalisation of Accounting Standards* deals with the role of the International Accounting Standards Board (IASB) and International Financial Reporting Standards (IFRS) in the economic world, but I am pleased to see that it also considers the efforts and roles of other standard setters. As a whole, the book tries to achieve a deeper understanding of the current situation by helping to provide answers to some key questions: Why is the accounting world following IFRS? Why do different countries follow different approaches to the globalization of accounting? Is it possible to have a common language if there are different interpretations?

The editors, Jayne Godfrey and Keryn Chalmers, mention in the introduction that "This book is born of a curiosity as to why countries would relinquish their existing national accounting standard-setting regimes to join the global movement" (p. 2). To satisfy this curiosity they have included 15 papers written by different authors. The wide representation of authors in terms of nationalities and backgrounds helps to guarantee a broad variety of approaches to the analysis of the globalization of accounting, its causes and consequences. As an example of the advantages of the diversity of views, it is useful to consider the different opinions on the proposal by the European Union (EU) in 2000 to adopt IFRS, which were then known as International Accounting Standards (IAS). Kevin Stevenson adopts a political view and argues that some probably felt they could influence the decisions of the International Accounting Standards Committee (IASC). He maintains that IAS could be seen as a softer alternative to U.S. GAAP, but still convenient (p. 35); similarly, David Alexander considers that the decision to follow IAS was made in order to avoid EU companies using U.S. GAAP, "over which Europe would have no influence whatever" (p. 52). However, Ruth Picker (p. 100) takes an economic view and argues that this was a direct response to the call to develop a single, more transparent European capital market. Interestingly, she states that the announcement made in July 2002 by the Chairman of the Australian Financial Reporting Council about the adoption of IFRS was a reaction to the EU decision made some days earlier. The Council of Ministers approved the IAS regulation on the 7th of June (pp. 98–99). There is no doubt that each of these individual views constituted only some of the many arguments that were raised before the EU between 2000, when the European Commission made its proposal, and 2002, when the EU approved the regulation, and one should point out that it was in the middle of this two-year period that the part-time IASC was transformed into the current independent, full-time body, the IASB.

Although the editors do not explicitly mention who the intended audience is, anyone interested in accounting information is clearly a potential reader of the book. Thus, we should look at the entire spectrum of stakeholders as interested readers. Indeed, in the foreword Sir David Tweedie commends the book to all readers seeking insights into the depth and breadth of issues relating to the globalization of accounting standards.

The book includes an introductory chapter written by the editors followed by another 15 previously unpublished papers, which in my view enrich the text and provide a fresh perspective, although some details, once printed, are quickly outdated due to the march of events, such as the decision by the U.S. Securities and Exchange Commission (SEC) in November 2007 to eliminate the required U.S. GAAP reconciliation for IFRS users. The next two chapters offer a more general content. In the first, Gordon Clark, Tessa Hebb, and Dariusz Wójcik deal with accounting as a language and discuss the influence of the Anglo-American world in continental Europe in terms of institutions and language. The second paper by Stevenson describes the role of the IASB (as well as the IASC) in the accounting world. Throughout the chapter he insists on something that might appear obvious: the user focus of IFRS. But, as he explains, this is not so well understood by some stakeholders. The remaining chapters focus on selected countries, and although the selection criterion follows a certain rationale, explained in the introduction, it will probably not convince all readers.

Chapters 4 to 7 are written by authors from the four countries whose standard setters formed the G4 (G4+1 if the IASC is considered): the U.K., the U.S., Canada, and Australia. Alexander uses the Darwinian principle to explore the British influence on the globalization of accounting standards. He starts with the EU's Accounting Directives issued in the last quarter of the past century and the introduction of the "true and fair view" notion, and continues with the IASC, whose accounting model embraces the Anglo-American world. Donna Street explains the role of the Financial Accounting Standards Board (FASB) in the globalization of standards since 1991, when it modified its domestic perspective and invested a lot of effort into the formation and development of the G4+1. She also refers to current concerns, such as the Memorandum of Understanding toward convergence between the IASB and the FASB, and the SEC's roadmap to eliminate reconciliations for foreign registrants using IFRS. Chapters 6 and 7 refer to Canada and Australia, respectively. James Gaa discusses the "principles versus rules" approach as the two views have influenced the Canadian accounting model. Picker provides an interesting analysis of the Australian experience where, due to the approach followed by the AASB (Australian Accounting Standards Board) of having a single set of sector-neutral accounting standards, both public-sector and unlisted reporting entities have also been obliged to use IFRS. This is further discussed in depth in a later chapter by Chalmers, Godfrey, Ian Langfield-Smith, and Wei Lu, where they explain that this decision forced the AASB to introduce specific paragraphs into the IFRS to render them appropriate for not-for-profit sector entities; consequently the so-called "Australian equivalents of IFRS" differ from IFRS. They also discuss the efficiency and efficacy of such a decision given the parallel development of IFRS and IPSAS (International Public Sector Accounting Standards) issued by the IPSASB (International Public Sector Accounting Standards Board).

The next three chapters were written by representatives of EU countries: Italy, France, and Germany. As in other EU countries, listed companies have to follow endorsed IFRS in their consolidated financial statements. Nevertheless, the EU's IAS regulation of 2002 also allows countries to require or permit the use of these standards for private companies and parent company financial statements. Consequently, they might have an even wider impact, as indeed some countries are revising their legislation to require IFRS more broadly, while others have allowed companies to use them.[1] As recognized by Angelo Provasoli, Pietro Mazzola, and Lorenzo Pozza, despite having a board member on the IASC, Italy played a marginal role in the international standard-setting context. The accounting system has traditionally been civil law-based, tax-dominated, and conservatively oriented, characteristics that do not fit very well with IFRS. Nevertheless, since 1990 there have been impressive changes in the Italian socio-economic context, and the new accounting body, OIC (*Organismo Italiano di Contabilità*), created in 2001, has required the adoption of IFRS for the parent company financial statements of listed companies, too. To facilitate the adoption, OIC will issue guidance taking interpretations of the IASB's International Financial Reporting Interpretations Committee (IFRIC) into consideration. This is particularly critical, because parent company financial statements are used for tax and legal purposes, as in some other European countries. Because of these links, in the following chapter by Serge Evraert and Jean-François des Robert, the authors make the point that French companies are worried about the impact of convergence with IFRS on taxation; they foresee an increase in the tax burden due to a reduction in the number of accepted provisions and the application of the fair value criterion for some assets. In fact, they recognize that there are numerous discrepancies between IFRS and the French accounting principles from a philosophical, methodological, and practical perspective (p. 135). As a result, one can imagine a dual accounting model where IFRS and local GAAP are used for different purposes, although there is a progressive adaptation of the law permitting the convergence to IFRS. The chapter by Hans Peter Möller follows a completely different approach. He analyzes empirically the superiority of U.S. GAAP and IFRS versus German GAAP. Nevertheless, his results do not allow confirmation of consistent differences in risk profiles based on the accounting model.

Three of the next papers deal with countries where IFRS are probably the ultimate goal, but do not intend to follow them directly: China, Japan, and India. Each country has its own peculiarity; so, as the authors argue, a direct application of IFRS does not seem reasonable at present, but in the long term they are committed to a convergence process. Nevertheless, as Wei-Guo Zhang and De-Ming Lu argue, "How to converge with IFRS is a matter for China to determine" (p. 211). Chitoshi Koga and Gunnar Rimmel argue that in Japan the concept of decision usefulness is peripheral to the stewardship function. Although, since 2005, there has been a convergence program between the IASB and the ASBJ (Accounting Standards Board of Japan), the results of a questionnaire sent to the 500 largest companies were not entirely satisfactory: Seventy percent of the respondents thought the costs of adopting IFRS were likely to exceed the benefits. As R. Narayanaswamy indicates, India has also seen recent significant changes, mainly due to the pressure from IOSCO (International Organization of Securities Commissions) and the SEBI (Securities and Exchange Board of India), and the standards published after 2000 follow IFRS more closely than previously. However, the author admits that there are many differences, and in many cases only the benchmark treatment included in IFRS is allowed by Indian standards.

The two last papers take a different view of the impact of accounting standards. The former, by Norita Mohd Nasir and Aniza Zainol, refers to Malaysia and discusses the role that religion can play in accounting standard setting. As long as IFRS are based on user needs and in particular on the needs of capital market users, it appears doubtful that they can be useful for a context where concepts of interest and profit do not exist. The latter paper, by Iain Edwards, Peter Schelluch, Adel Du Plessis, Jean Struweg, and Andrew West, argues that the decision to adopt IFRS by the South African government has been critical for the economic development of the country. Nevertheless, in common with some other chapters, they urge the further education and training of the accounting profession in order that IFRS might be applied properly.

Overall, the book is a stimulating contribution to the debate on the globalization of accounting. The authors analyze the topic from different angles and perspectives, and it is the reader's task to discover answers to key questions (some mentioned above). The book is "relevant" (as this term is defined in the IASB Framework) because it enables the readers to respond to the aforementioned questions. Thus, I would highly recommend it to those interested in recent accounting changes, but perhaps it will prove more suited to scholars and policy makers. In my view, to know about the past is a useful way to learn about (and guide) the future changes that inevitably will occur due to the ever-changing environment in which we live.

---

[1] http://ec.europa.eu/internal_market/accounting/ias_en.htm#options.

BEGOÑA GINER
Professor of Accounting and Finance
Universitat de València

NICOLAS VÉRON, *The Global Accounting Experiment* (Brussels, Belgium: Bruegel Blueprint Series, Volume II, 2007, ISBN 978-9-078910-02-2, pp. vii, 71, €10 but also available online free of charge at http://www.bruegel.org).

The author is a research fellow at Bruegel and also a business columnist with a practical background in public policy and corporate finance. Bruegel is a European think tank financed by European governments and international corporations with the aim of contributing to the quality of economic policymaking in Europe. This booklet aims at an audience of policymakers rather than accounting academics. For the latter group, the booklet is an interesting example of secondary literature produced to influence the political process.

The "global accounting experiment" is the decision by the European Union (EU) to adopt International Financial Reporting Standards (IFRS) for listed companies and the decisions of other jurisdictions (e.g., China, Australia, and Canada) also to adopt IFRS. The motivation for the booklet is a fear over the sustainability of this development. Two critical issues are highlighted: the problem of the legitimacy of the International Accounting Standards Board (IASB), and the question of consistent IFRS implementation in the various jurisdictions. In the last chapter, the author develops "recommendations for sustainability."

The booklet starts with a short history of accounting and standard setting (Chapter 1, pp. 5–25), which starts in the medieval times and then rushes to the development of the current institutions of standard setting (e.g., the Financial Accounting Standards Board [FASB] and the IASB), finally ending with the EU's adoption of IFRS. In Chapter 2 (pp. 26–33), the author describes "the forces at play" in the development of the former International Accounting Standards Committee (IASC), emphasizing in particular the role of the U.S.-based constituencies. It should be added that European constituencies were also very supportive of the formation of the IASC in the 1970s, when at the same time a first major effort to harmonize accounting regulation in the EU was undertaken, which led to the Fourth Directive on individual company financial statements in 1978 and the Seventh Directive on consolidated financial statements in 1983, followed by the Eighth Directive on auditing. The implementation of those directives in the national law of the EU member countries took place during the 1980s and early 1990s and resulted in major changes in their national accounting regulations. Through the directives, the EU exercised the power to regulate financial accounting based on the 1957 Treaty of Rome. I strongly disagree with the statement by the author that "standard-setting remained a national task which was only marginally impacted by the [Fourth and Seventh] directives" (p. 30). The author is right, however, in that the EU harmonization process stopped in the 1980s as the EU failed to respond to the needs of the large European multinational companies for a globally accepted set of accounting standards.

Beginning in the middle 1990s, many of the leading European multinational companies, in particular from the German language countries, switched to International Accounting Standards (IAS) or to U.S. GAAP. This development was in part driven by stock exchange regulations requiring financial statements prepared by the use of internationally accepted accounting principles, especially for segments of some European stock exchanges catering to small, high-growth companies as well as because of options granted by national accounting laws to enable the use of IAS or U.S. GAAP in the consolidated financial statements of listed companies (see Daske and Gebhardt 2006). In my view, the author neglects two important developments. The first was the huge effort by the IASC to improve IAS in order to meet the requirements for being accepted by the International Organization of Securities Commissions (IOSCO), resulting in major improvements of IAS by the end of 1998. The second was the restructuring of the IASC to become the IASB by 2001. It should be noted that the support and/or pressure of the U.S. Securities and Exchange Commission (SEC) and the FASB were instrumental in shaping the IASB in its current form. It is inconceivable that the EU regulation to adopt IFRS would have been passed without the restructuring of the IASB and the improvements in its standards by 1998.

In Chapter 3, "The Next Challenges" (pp. 34–47), the author elaborates on the legitimacy challenge. It is a fact that the IASB has no democratically derived power to prescribe the application of accounting standards. For the EU's member countries, this power has been delegated to the EU, which adopts the IASB's standards via a special endorsement process. This makes, in principle, the IASB vulnerable to pressure from the EU constituency. However, the IASB has demonstrated its relative independence in the IAS 39 carve-out controversy. Not endorsing

IASB standards is not really an option for the EU in the absence of an alternative to the IASB for accounting standard setting. The author points to the pressure from U.S. constituencies, in particular the SEC and FASB, and—in his view—the detrimental effects of the convergence process on the quality of the accounting standards. As an example, he points to the problem of IFRS 8 that is essentially a copy of the FASB's FAS 131 and is characterized by the author as being of lower quality than the previous IAS 14 standard. Here and elsewhere in the booklet, the author suggests to the reader that the quality of standards is readily measurable and that there is no dispute about such measurements. Unfortunately, this is a most tricky issue—much recent research has reached diverse findings and rarely with conclusive results.

Another major concern of the author is with the funding of the IASB, which is the responsibility of the trustees of the IASC Foundation, who in turn select the IASB Board members. Being dependent on voluntary contributions makes the IASB vulnerable to the donors. In Chapter 4 (pp. 48–63), on recommendations, these concerns are addressed with a very general recommendation to make the trustees accountable and to make funding sustainable. How this reasonably could be achieved is not further explored.

Another proposal is to "Put quality above speed in IFRS/US GAAP convergence" (pp. 51–52). The key issue of how to measure quality is not addressed. The author proposes that the EU "exert legitimate and constructive pressure on the IASB by not endorsing standards" (p. 55) and proposes to try this on the case of IFRS 8 on segmental reporting. As I write this review, the European Parliament has finally voted to endorse IFRS 8. Further, the author recommends to the European Commission that it should "call for an analysis of past and current reporting under, respectively, IAS 14 and its FAS 131 counterpart" (p. 56). In fact, the European Commission solicited several research studies on the introduction of IFRS in 2005 (e.g., ICAEW 2007) and is sponsoring an EU-wide research program INTACCT, "The European IFRS Revolution: Compliance, Consequences and Policy Lessons," under the Commission's 6th framework Marie Curie program (http://www.intacct-research.org). Here the author misses the opportunity to make a constructive proposal to supplement the IASB's standard setting with theoretical and empirical research on the economic effects of standard setting.

A recommendation to the U.S. is to accept the IASB's global promise and respect the autonomy of the IASB as a standard setter. With the recent SEC decision to eliminate the reconciliation requirement for foreign companies listed in the U.S. that use IFRS, the position of the IASB toward pressures from the EU or other national constituencies, including the U.S., has been considerably strengthened.

Another challenge is that of the consistent implementation of IFRS. This challenge is addressed by a recommendation to set up an accounting authority for the EU to provide guidance for IFRS implementation at the EU level. However, this creates the risk of national or regional interpretations of IFRS that is not consistent with the ultimate goal of one global set of high quality accounting rules—a goal that is explicitly shared by the author (p. 4).

To summarize, the recommendations by the author are not rigorously derived but are more or less policy statements for an EU-geocentric target group of policymakers.

The bibliography (pp. 66–69) is far from being comprehensive in covering the literature on the effects of adopting IFRS. The language of the booklet is persuasive and may well have the intended impact on the target group of policymakers. For accounting academics, the booklet is worthwhile to read as an example how professional think tanks contribute to regulatory processes.

REFERENCES

Daske, H., and G. Gebhardt. 2006. International financial reporting standards and experts' perceptions of disclosure quality. *ABACUS* 42 (3/4): 461–498.

The Institute of Chartered Accountants in England and Wales (ICAEW). 2007. *EU Implementation of IFRS and the Fair Value Directive—A Report for the European Commission.* London, U.K.: The Institute of Chartered Accountants in England and Wales (available at: http://www.icaew.com/index.cfm?route=145392).

GÜNTHER GEBHARDT
Professor of Accounting and Auditing
Johann Wolfgang Goethe-Universität Frankfurt am Main

CAPSULE COMMENTARY

By the Book Review Editor

GIOVANNI FRATTINI (editor), *Improving Business Reporting: New Rules, New Opportunities, New Trends* (Milano, Italy: Giuffrè Editore, 2007, ISBN 88-14-13580-0, pp. xvi, 832).

This is the proceedings volume from the 8th International Conference held in Novara, Italy on November 24–25, 2006 to celebrate the work of young researchers. It was held under the auspices of the Accademia Italiana di Economia Aziendale with the cooperation of the Università degli Studi del Piemonte Orientale "Amedeo Avogadro." It contains 13 papers on accounting (mostly dealing with IFRS issues), eight on finance, and nine on management and corporate governance. They demonstrate an impressive awareness of the international literature. Of the 60 papers presented during the parallel sessions at the conference, a scientific committee chose the 30 to be reproduced in this volume. All of the papers are written in the English language.

Editorial Data

The following table contains information about turnaround time for manuscripts (including revisions) on which editorial decisions were made in the 12-month period ended May 31, 2007. Turnaround time is the number of days between the date that the manuscript was received *and the date of the editor's letter to the author(s).*

	Number of Manuscripts	Percent	Cumulative Number	Cumulative Percent
0 ≤ Days ≤ 30	39	8.32	39	8.32
31 ≤ Days ≤ 60	153	32.62	192	40.94
61 ≤ Days ≤ 90	153	32.62	345	73.56
91 ≤ Days ≤ 120	86	18.34	431	91.90
121 ≤ Days	38	8.10	469	100.00
Total	469	100.00		

The mean review time was 70; the median review time was 66 days.

New Submissions by calendar year (January 1 to December 31)

1992	241
1993	234
1994	231
1995	195
1996	230
1997	195
1998	196
1999	239
2000	260
2001	260*
2002	324
2003	327
2004	307
2005	389
2006	413
2007	443

* 2001—in addition to the 260 regular submissions, 68 MSs were submitted to the *TAR* Quality of Earnings Conference.

The acceptance rate (defined as number of MSs accepted divided by the number of new submissions) over the past 5 years is between 12–16% per year.

EDITORIAL POLICY AND STYLE INFORMATION

EDITORIAL POLICY

According to the policies set by the Publications Committee (which were endorsed by the Executive Committee and were published in the *Accounting Education News*, June 1987), *The Accounting Review* "should be viewed as the premier journal for publishing articles reporting the results of accounting research and explaining and illustrating related research methodology. The scope of acceptable articles should embrace any research methodology and any accounting-related subject, as long as the articles meet the standards established for publication in the journal ... No special sections should be necessary. The primary, but not exclusive, audience should be—as it is now—academicians, graduate students, and others interested in accounting research."

The primary criterion for publication in *The Accounting Review* is the significance of the contribution an article makes to the literature.

The efficiency and effectiveness of the editorial review process is critically dependent upon the actions of both authors submitting papers and the reviewers. Authors accept the responsibility of preparing research papers at a level suitable for evaluation by independent reviewers. Such preparation, therefore, should include subjecting the manuscript to critique by colleagues and others and revising it accordingly prior to submission. The review process is not to be used as a means of obtaining feedback at early stages of developing the research.

Reviewers and associate editors are responsible for providing critically constructive and prompt evaluations of submitted research papers based on the significance of their contribution and on the rigor of analysis and presentation. Associate editors also make editorial recommendations to the editor.

MANUSCRIPT PREPARATION AND STYLE

The Accounting Review's manuscript preparation guidelines follow *The Chicago Manual of Style* (15th ed.; University of Chicago Press). Another helpful guide to usage and style is *The Elements of Style*, by William Strunk, Jr., and E. B. White (Macmillan). Spelling follows *Webster's Collegiate Dictionary.*

FORMAT

1. All manuscripts should be printed in 12-point font on one side of 8½ × 11" good quality paper and be double-spaced, except for indented quotations.
2. Manuscripts should be as concise as the subject and research method permit, generally not to exceed 7,000 words.
3. Margins of at least one inch from top, bottom, and sides should facilitate editing and duplication.
4. To promote anonymous review, authors should not identify themselves directly or indirectly in their papers or in experimental test instruments included with the submission. Single authors should not use the editorial "we."
5. A cover page should show the title of the paper, the author's name, title and affiliation, email address, any acknowledgments, and a footnote indicating whether the author would be willing to share the data (see last paragraph in this statement).

Pagination: All pages, including tables, appendices and references, should be serially numbered. Major sections should be numbered in Roman numerals. Subsections should not be numbered.

Numbers: Spell out numbers from one to ten, except when used in tables and lists, and when used with mathematical, statistical, scientific, or technical units and quantities, such as distances, weights and measures. For example: *three days; 3 kilometers; 30 years.* All other numbers are expressed numerically.

Percentages and Decimal Fractions: In nontechnical copy use the word percent in the text.

Hyphens: Use a hyphen to join unit modifiers or to clarify usage. For example: *a well-presented analysis; re-form.* See *Webster's* for correct usage.

Keywords: The abstract is to be followed by four keywords that will assist in indexing the paper.

ABSTRACT/INTRODUCTION

An Abstract of about 100 words should be presented on a separate page immediately preceding the text. The Abstract should concisely inform the reader of the manuscript's topic, its methods, and its findings. Keywords and the Data Availability statements should follow the Abstract. The text of the paper should start with a section labeled "I. Introduction," which provides more details about the paper's purpose, motivation, methodology, and findings. Both the Abstract and the Introduction should be relatively nontechnical, yet clear enough for an informed reader to understand the manuscript's contribution. The manuscript's title, but neither the author's name nor other identification designations, should appear on the Abstract page.

TABLES AND FIGURES

The author should note the following general requirements:

1. Each table and figure (graphic) should appear on a separate page and should be placed at the end of the text. Each should bear an Arabic number and a complete title indicating the exact contents of the table or figure. Tables and figures should define each variable. The titles and definitions should be sufficiently detailed to enable the reader to interpret the tables and figures without reference to the text.
2. A reference to each graphic should be made in the text.
3. The author should indicate by marginal notation where each graphic should be inserted in the text.
4. Graphics should be reasonably interpreted without reference to the text.
5. Source lines and notes should be included as necessary.

Equations: Equations should be numbered in parentheses flush with the right-hand margin.

DOCUMENTATION

Citations: Within-text citations are made using an author-year format. Cited works must correspond to the list of works listed in the "References" section. Authors should make an effort to include the relevant page numbers in the within-text citations.

1. In the text, works are cited as follows: author's last name and year, without comma, in parentheses. For example: one author, (Berry 2003); two authors, (Fehr and Schmidt 2003); three or more authors, (Dechow et al. 1998); more than one work cited, (Cole and Yakushiji 1984; Dechow et al. 1995; Levitt 1998); with two works by the same author(s), (Nelson 2003, 2005).
2. Unless confusion would result, do not use "p." or "pp." before page numbers. For example, (Dechow and Dichev 2002, 41–42).
3. For cited works that include more than one work by an author (or same co-authors) that are published in the same year, the suffix a, b, etc., is to follow the date in the within-text citations and in the "References" section. For example: (Baiman and Rajan 2002a, 2002b).
4. When the author's name is mentioned in the text, it need not be repeated in the citation. For example: "Cohen et al. (2005) provide ..."
5. Citations to institutional works should use acronyms or short titles where practicable. For example: (NCFFR, The Treadway Commission 1987).
6. If the paper refers to statutes, legal treatises, or court cases, citations acceptable in law reviews should be used.

Reference List: Every manuscript must include a "References" section that contains only those works cited within the text. Each entry should contain all information necessary for unambiguous identification of the published work. Use the following formats (which follow *The Chicago Manual of Style*):

1. Arrange citations in *alphabetical order* according to the surname of the first author or the name of the institution or body responsible for the published work.
2. Arrange citations in *chronological order* when two or more works are by the same first author (regardless of co-authors). Two or more works by the *same author(s)* published in the same year are distinguished by letters a, b, etc., after the year.
3. Use authors' initials instead of proper names.
4. For two or more authors, separate authors with a comma, including a comma before "and" (Dechow, P. M., R. Sloan, and A. Sweeney).
5. Date of publication follows the author's (authors') name(s).
6. Titles of journals or newspapers should not be abbreviated.

Sample entries are as follows:

Baiman, S., and M. Rajan. 2002a. The role of information and opportunism in the choice of buyer-supplier relationships. *Journal of Accounting Research* 40 (2): 247–278.

———, and ———. 2002b. Incentive issues in inter-firm relationships. *Accounting, Organizations and Society* 27 (3): 213–238.

Berry, R. 2003. Testimony before the Senate Committee on Homeland Security and Governmental Affairs Permanent Subcommittee on Investigations. November 18. Available at: http://hsgac.senate.gov/_files/111803berry.pdf.

Cohen, D., A. Dey, and T. Lys. 2005. The Sarbanes Oxley Act of 2002: Implications for compensation structure and risk-taking incentives of CEOs. Working paper, New York University, University of Chicago, and Northwestern University.

Cole, R., and T. Yakushiji, eds. 1984. *The American and Japanese Auto Industries in Transition.* Ann Arbor, MI: University of Michigan.

Dechow, P. M., R. Sloan, and A. Sweeney. 1995. Detecting earnings management. *The Accounting Review* 70 (2): 193–225.

———, S. P. Kothari, and R. L. Watts. 1998. The relation between earnings and cash flows. *Journal of Accounting and Economics* 25: 133–168.

———, and I. Dichev. 2002. The quality of accruals and earnings: The role of accrual estimation errors, *The Accounting Review* 77 (Supplement): 35–59.

Easton, P. 2003. Discussion of: The predictive value of expenses excluded from pro forma earnings. *Review of Accounting Studies* 8: 175–183.

Engel, E., R. Hayes, and X. Wang. 2007. The Sarbanes-Oxley Act and firms' going-private decisions. *Journal of Accounting and Economics* (forthcoming).

Financial Accounting Standards Board (FASB). 2006. *Accounting for Uncertainty in Income Taxes, and Interpretation of FASB Statement No. 109.* FASB Interpretation No. 48. Financial Accounting Series. Norwalk, CT: FASB.

Fehr, E., and K. Schmidt. 2003. A theory of fairness, competition, and cooperation. In *Advances in Behavioral Economics*, edited by C. Camerer, G. Loewenstein, and M. Rabin, 271–296. New York, NY: Princeton University Press.

Janis, I. L. 1982. *Groupthink: Psychological Studies of Policy Decisions and Fiascoes.* Boston, MA: Houghton Mifflin.

Levitt, A. 1998. The numbers game. Speech delivered at New York University, Center for Law and Business, September 28.

Maggi, G. 1999. The value of commitment with imperfect observability and private information. *RAND Journal of Economics* (Winter) 30: 555–574.

National Commission on Fraudulent Reporting (the Treadway Commission). 1987. *Report of the National Commission on Fraudulent Financial Reporting.* Washington, D.C.: NCFFR.

Nelson, M. W. 2003. Behavioral evidence on the effects of principles- and rules-based standards. *Accounting Horizons* 17 (1): 91–104.

———. 2005. A review of experimental and archival conflicts-of-interest research in auditing. In *Conflicts of Interest: Challenges and Solutions in Business, Law, Medicine, and Public Policy,* edited by D. A. Moore, D. M. Cain, G. Loewenstein, and M. H. Bazerman. Cambridge, U.K.: Cambridge University Press.

Rigdon, E. E., R. E. Schumacker, and W. Wothke. 1998. A comparative review of interaction and nonlinear modeling. In *Interaction and Nonlinear Effects in Structural Equation Modeling*, edited by R. E. Schumacker, and G. A. Marcoulides, 1–16. Mahwah, NJ: Erlbaum Associates.

Schultz, E., and T. Francis. 2002. Companies profit on workers' deaths through "dead peasants" insurance. *Wall Street Journal* (April 19): 1.

Securities and Exchange Commission (SEC). 2002. *Certification of Disclosure in Companies' Quarterly and Annual Reports.* Release Nos. 33-8124, 34-46427. Washington, D.C.: SEC.

U.S. House of Representatives. 2002. The Sarbanes-Oxley Act of 2002. Public Law 107-204 [H. R. 3763]. Washington, D.C.: Government Printing Office.

Footnotes: Footnotes are not used for documentation. Textual footnotes should be used only for extensions and useful excursions of information that if included in the body of the text might disrupt its continuity. Footnotes should be consecutively numbered throughout the manuscript with superscript Arabic numerals. Footnote text should be doubled-spaced and placed at the end of the article.

SUBMISSION OF MANUSCRIPTS

Authors should note the following guidelines for submitting manuscripts:

1. Manuscripts currently under consideration by another journal or publisher should not be submitted. The author must state that the work is not submitted or published elsewhere.
2. For manuscripts reporting on field surveys or experiments: If the additional documentation (e.g., questionnaire, case, interview schedule) is sent as a separate file, then all information that might identify the author(s) must be deleted from the instrument.
3. MS should be submitted via email as Microsoft® Word or PDF files to the *TAR* office at TARUAZ@email.arizona.edu. A submission letter, a separate cover page, and the MS containing no author identification should be emailed as separate files. The submission fee should be paid online (VISA or MasterCard only) at: https://aaahq.org/AAAforms/journals/tarsubmit.cfm. If unable to pay online, a check in U.S. funds, made payable to the American Accounting Association, may be sent to Professor Dan Dhaliwal, Senior Editor, *The Accounting Review*, The University of Arizona, Department of Accounting, McClelland Hall 301, 1130 E. Helen Street, Tucson, AZ 85721-0108. The submission fee in U.S. funds is $125.00 for members and $200.00 for nonmembers of the AAA. Effective January 1990, the submission fee is nonrefundable.
4. The author should retain a copy of the paper.
5. Revisions must be submitted within 12 months from request, otherwise they will be considered new submissions.

COMMENTS

Comments on articles previously published in *The Accounting Review* will be reviewed (anonymously) by two reviewers in sequence. The first reviewer will be the author of the original article being subjected to critique. If substance permits, a suitably revised comment will be sent to a second reviewer to determine its publishability in *The Accounting Review*. If a comment is accepted for publication, the original author will be invited to reply. All other editorial requirements, as enumerated above, also apply to proposed comments.

POLICY ON REPRODUCTION

An objective of *The Accounting Review* is to promote the wide dissemination of the results of systematic scholarly inquiries into the broad field of accounting.

Permission is hereby granted to reproduce any of the contents of the *Review* for use in courses of instruction, as long as the source and American Accounting Association copyright are indicated in any such reproductions.

Written application must be made to the American Accounting Association, 5717 Bessie Drive, Sarasota, FL 34233-2399, for permission to reproduce any of the contents of the *Review* for use in other than courses of instruction—e.g., inclusion in books of readings or in any other publications intended for general distribution. In consideration for the grant of permission by the *Review* in such instances, the applicant must notify the author(s) in writing of the intended use to be made of each reproduction. Normally, the *Review* will not assess a charge for the waiver of copyright.

Except where otherwise noted in articles, the copyright interest has been transferred to the American Accounting Association. Where the author(s) has (have) not transferred the copyright to the Association, applicants must seek permission to reproduce (for all purposes) directly from the author(s).

POLICY ON DATA AVAILABILITY

The following policy has been adopted by the Executive Committee in its April 1989 meeting. "An objective of (*The Accounting Review, Accounting Horizons, Issues in Accounting Education*) is to provide the widest possible dissemination of knowledge based on systematic scholarly inquiries into accounting as a field of professional research, and educational activity. As part of this process, authors are encouraged to make their data available for use by others in extending or replicating results reported in their articles. Authors of articles which report data dependent results should footnote the status of data availability and, when pertinent, this should be accompanied by information on how the data may be obtained."

MAY 2008 PLACEMENT ADS

*The deadline for information to be included in the "Placement Information" section is two months prior to the date of publication. The **Review** is published in January, March, May, July, September, and November. Placement advertising and replies should be sent to the American Accounting Association, 5717 Bessie Drive, Sarasota, FL 34233-2399, ATTENTION: ADVERTISING DEPT. All replies to box numbers should be on separate sheets to facilitate forwarding.*

301 UNIVERSITY OF NORTH TEXAS invites applications for two or more tenure-track faculty positions effective Fall 2008. Areas of need include: Auditing, Cost/Managerial, Financial, Information Systems, and Taxation. The successful candidate must have an earned doctorate in Accounting by August 2008 and a commitment to excellence in research and teaching. Applications for Assistant, Associate, or Professor will be considered. A national reputation based on record and recent publications in top-quality journals is expected for senior rank candidates. Professional certification is preferred. Salaries are competitive. UNT is a comprehensive university, offering undergraduate, master's, and doctoral programs. The College of Business Administration is accredited by AACSB and the Department of Accounting has separate AACSB accreditation. More information about the department is available at: http://www.coba.unt.edu/acct/. The University of North Texas is an emerging research institution located in the vibrant and rapidly expanding Dallas-Fort Worth metropolitan area and situated in close proximity to DFW International Airport. University enrollment is over 34,000 students with more than 6,500 graduate students. UNT offers 257 undergraduate and graduate majors and concentrations. Review of applications will begin immediately and continue until the positions are filled. Please mail letter of interest and *vita* with references to: Dr. Paul D. Hutchison, Department of Accounting, PO Box 305219, Denton, TX 76203-5219, or send electronic submissions to: naudei@unt.edu. The University of North Texas is an Equal Opportunity/Affirmative Action Employer.

302 MASSACHUSETTS INSTITUTE OF TECHNOLOGY, Alfred P. Sloan School of Management, MIT Sloan School's Accounting Group seeks new faculty at all levels (Assistant, Associate, and Full Professor) for the academic year beginning July 2008, or as soon as possible thereafter. Applicants will be evaluated on their potential for and/or record of excellence in research and teaching. Please include evidence of research such as a publication, working paper, or dissertation proposal with other application materials. The deadline for applications is March 1, 2008. However, applications will be evaluated as soon as they are complete, if submitted at an earlier date. Applications must include a cover letter, *curriculum vitae*, and letters of recommendation. Applications should be sent to: Professor Joseph Weber, Chair, Accounting Group Search Committee, MIT Sloan School of Management, 50 Memorial Drive, E52-325, Cambridge, MA 02142-1347. MIT is an Equal Opportunity Employer committed to building a culturally diverse intellectual community and strongly encourages applications from women and minorities.

303 UNIVERSITY OF WISCONSIN–PARKSIDE, School of Business and Technology invites applications for a full-time position as an Assistant/Associate Professor of Accounting beginning August 2008. The applicants should have a Ph.D. or D.B.A. in Accounting. ABD will be considered. Applicants must be able to teach managerial accounting at both undergraduate and M.B.A. levels. Though not required, a second area of teaching interest in auditing would be a plus. Applicants must provide evidence of quality teaching and a commitment to scholarly research. Interested individuals should send letter of application, current *curriculum vitae*, transcripts (unofficial copies acceptable), recent teaching evaluations, examples of scholarly work, and the names, telephone numbers, and email addresses of three professional references to: Professor Jamie Wang, School of Business & Technology, University of Wisconsin–Parkside, 900 Wood Rd., PO Box 2000, Kenosha, WI 53141-2000. For more information, see the following link: http://newweb.uwp.edu/departments/human.resources/unclassified.positions/acct.cfm. UWP is an Equal Opportunity Employer.

304 WESTERN NEW ENGLAND COLLEGE invites applications for two tenure-track Assistant/Associate Professor accounting positions to begin August 2008. Terminal degree is required by starting date. Successful candidates will demonstrate a strong commitment to teaching financial accounting, auditing, or accounting information systems/forensic accounting in both our undergraduate and graduate programs, and to scholarly research and journal publication. Western New England College is a private, independent, coeducational institution located on an attractive 215-acre suburban campus in Springfield, Massachusetts within easy access of Boston and New York City. The School of Business is accredited by AACSB International. More information about the College is available at: http://www.wnec.edu. Send *vita*, teaching evaluations, research manuscript, and reference information to: Dr. Julie Siciliano, Dean, School of Business, Western New England College, 1215 Wilbraham Road, Springfield, MA 01119. Electronic submissions are encouraged and should be sent to Angie Jacque at: ajacque@wnec.edu. Western New England College is an Equal Opportunity Employer.

305 PORTLAND STATE UNIVERSITY, School of Business Administration invites applications for a full-time Accounting faculty member. This one-year fixed term appointment (September 15, 2008–June 15, 2009) may be renewed upon mutual agreement. A Ph.D. or equivalent in Accounting is preferred. Applicants with primary emphasis on financial accounting or Accounting Information Systems given preference. Candidates must contribute to the Accounting Area's mission in the areas of teaching, curriculum development, and service. The candidate must be able to deliver excellent undergraduate instruction and have an appreciation for the role of accounting in society. Salary is competitive, dependent upon qualifications and experience. Applicants should provide a letter of interest, *curriculum vita*, evidence of teaching effectiveness, and names of three references to: Dr. Darrell Brown, School of Business Administration, Portland State University, PO Box 751, Portland, OR 97207-0751; Phone: (503)725-3096; Email: darrellb@sba.pdx.edu. Review of applications begins immediately and is open until finalists are identified.

306 TEXAS A&M UNIVERSITY–CORPUS CHRISTI, Department of Accounting and Business Law, College of Business, invites applications for a tenure-track Assistant/Associate Professor position beginning Fall 2008, subject to final budgetary approval. The position is open to all areas of accounting specialization, with a preference for applicants in accounting information systems. Applicants should possess a doctoral degree in accounting from an AACSB-accredited institution and demonstrate a strong commitment to scholarly research and teaching excellence. Certification and professional experience are preferred. Salary is competitive. Please send applications to: Donald R. Deis, Chair, Accounting & Business Law Department, College of Business, Texas A&M University–Corpus Christi, 6300 Ocean Drive, Unit # 5808, Corpus Christi, Texas 78412-5808; or email: Donald.Deis@tamucc.edu. For more information on the positions, the University, and the College go to: http://www.cob.tamucc.edu. The College of Business undergraduate and graduate business programs are accredited by AACSB International. Gloria Ann Valdez, Business Coordinator Texas A&M University–Corpus Christi, College of Business, FC 128, 6300 Ocean Drive, Unit 5808, Corpus Christi, TX 78412-5808; Phone: (361) 825-5828; Fax (361) 825-2725; Email: gloria.valdez@tamucc.edu. TAMUCC is an AA/EEO employer committed to excellence (http://www.tamucc.edu/~hrweb/).

307 POSITION WANTED Full-time lecturer position desired by CPA/M.Tax. to start Summer or Fall 2008. Experience teaching financial accounting and taxation at university and community college levels. Outstanding teaching evaluations and references. Big 4 tax experience. Prefer a two-year position while I prepare Ph.D. applications. Prefer West or Northeast but will consider any location. Contact jstekelberg@yahoo.com.

308 OAKLAND UNIVERSITY AACSB-Accounting-accredited program invites applications for a tenure-track position in Accounting, beginning August 15, 2008. An interest in audit or accounting information systems is desirable but not essential. Scholarly work is supported and expected. Full details about the position and application procedures are available at: http://www2.oakland.edu/provost/web/jobdetail.cfm?ID=326. Oakland University is an Equal Opportunity Employer.

309 THE UNIVERSITY OF TEXAS AT SAN ANTONIO, Department of Accounting invites applications and nominations for the Glenn L. and Sally P. Ramsdell Memorial Chair for Accounting. Rank is at the Professor level beginning August 2008, pending budget approval. The Ramsdell Chair incumbent will be a research leader in the Department and a major contributor to the Ph.D. program. Applicants must submit a letter of application, evidence of teaching and research excellence, *vita*, and the names and contact addresses of three references to: Dr. Marshall Pitman, Search Committee Chair, at: marshall.pitman@utsa.edu. Applicants selected for interviews must be able to show proof that they are eligible and qualified to work in the United States. For detailed requirements consult the full text of the position announcement at: http://business.utsa.edu/jobs (or email Dr. Pitman). UTSA is an Affirmative Action/Equal Employment Opportunity Employer. Women, minorities, veterans and individuals with disabilities are encouraged to apply.

310 SUNGKYUNKWAN UNIVERSITY, SKK Graduate School of Business seeks applications for a tenure-track position (Assistant or Associate Professor) in Financial Accounting to begin in Fall 2008. SKK GSB is a business school newly founded by Sungkyunkwan University with the assistance of the Samsung Group and in collaboration with the MIT Sloan School of Management. Our mission is to become the best M.B.A. program in Asia. Salary and teaching loads are highly competitive. Generous fringe benefits include free housing, education allowance for international schools, and medical insurance. All candidates must have a doctoral degree in Accounting prior to the appointment. Candidates should show their strong potential for and/or record of excellence in research and M.B.A. teaching. Excellent English communication skills are required. Applicants should submit a cover letter, an updated *curriculum vitae*, an example of written work for presentation, three letters of reference, and evidence of teaching effectiveness to: Dr. Jae Ha Lee, Associate Dean, SKK GSB, Sungkyunkwan University, 53 Myungryun-dong 3-ga, Jongro-gu, Seoul, Korea 110-745; Phone: 82-2-740-1505, Email: gsbfaculty@skku.edu. Review of applications will begin immediately and continue until the position is filled. Additional information about SKK GSB can be found at: http://gsb.skku.edu.

311 UNIVERSITY OF SOUTH ALABAMA, Accounting Department invites applications for two tenure-track positions to teach undergraduate and graduate courses in accounting beginning Fall 2009. Applicants at all ranks will be considered. Applicants must possess a Ph.D. (or D.B.A) with a concentration in accounting or be ABD. Preferred qualifications include a primary teaching interest in auditing and/or financial accounting, professional certification (CPA and/or CIA), and an interest in professional service. The successful candidate should demonstrate excellence in teaching and a commitment to intellectual contributions. Applicants must submit a letter of application, current *curriculum vita*, and names (with phone numbers) of three references to: Dr. Russell Hardin, Chairperson, Department of Accounting, Mitchell College of Business, 307 University Blvd, Mobile, AL 36688. Electronic applications may be sent to: rhardin@usouthal.edu. Review of applications will begin March 17, 2008, and continue until the positions are filled. The University of South Alabama is an Equal Opportunity/Equal Access Employer.

312 UNIVERSITY OF NEVADA–RENO, Department of Accounting and Information Systems is seeking candidates for two full-time, tenure-track Assistant Professors in Accounting. The position is open to candidates in all accounting specialties. Duties of the position include teaching responsibilities at both the undergraduate and graduate (master's degree) level, intellectual contributions, and service activities within the university, academic, and professional communities. Candidates should possess a doctoral degree from an accredited program or be close to completion. For complete position description and requirements, contact: Search Chair/Coordinator, Dr. John Mills by email: jmills@unr.edu; or in writing: Department of ACC/IS, Mailstop 026, University of Nevada, Reno, Reno, NV 89557; or view at: http://jobs.unr.edu. Dr. Mills and members of the search committee will be attending the AAA Annual Meeting in Anaheim in August 2008. All interested candidates can apply directly with an online application or schedule a meeting at the AAA Annual Meeting to learn more about at the University of Nevada–Reno. These positions are open until filled. EEO/AA women and under-represented groups are encouraged to apply.

313 CANISIUS COLLEGE Accounting, The Wehle School of Business invites applicants at all levels for one tenure-track position in Accounting. Applicants for the position must possess a Ph.D. in accounting or be near completion of their dissertation. Areas of interest should be Auditing, AIS, and/or Forensics. The Department of Accounting seeks faculty with a strong commitment to teaching excellence balanced with an active interest in scholarly research to join its nationally recognized program. Canisius offers both graduate and undergraduate accounting degrees. The Wehle School of Business is accredited by the AACSB International. Send letters of interest and *curriculum vitae* to: Dr. Ian J. Redpath, Canisius College, 2001 Main St., Buffalo, NY 14208-1098. Canisius College is an independent, co-educational, medium-sized institution of higher education conducted in the Catholic and Jesuit tradition. Canisius College is an Equal Opportunity Employer. Minority candidates strongly encouraged to apply. Visit: http://www.canisius.edu/accjob.

314 MISSISSIPPI STATE UNIVERSITY, College of Business and Industry, Adkerson School of Accountancy invites applicants to apply for a Fall 2008 or Fall 2009 Endowed Chair (Full Professor) or Endowed Professorship (Associate or Full Professor) tenure-track position to assist with leadership and involvement in the School's doctoral program. Any area of specialization will be considered. Candidates must have an accounting doctorate degree, demonstrated excellence in teaching, publication record commensurate to that of an endowed chair or professorship, and experience in working with doctoral students. Professional certification and accounting experience are highly desirable. Compensation is competitive according to qualifications and experience. Mississippi State University, located in Starkville, MS, offers a unique university town setting with a great environment to live and work. Send letter of interest, current *curriculum vitae*, and names (including contact information) of three references to: Chair, Search Committee, Adkerson School of Accountancy, PO Box EF, Mississippi State, Starkville, MS 39762; or email Clyde Herring, Interim Director at: ceh4@msstate.edu. Position will remain open until filled. MSU is an AA/EEO Employer.

315 NANYANG TECHNOLOGICAL UNIVERSITY, Division of Accounting in the Nanyang Business School (NBS) invites applications for tenure-track faculty positions at all ranks in Accounting. All research areas will be considered. Visiting positions are also available. Job qualifications: Applicants should possess a Ph.D. in accounting, demonstrate or show promise of scholarly research publications in top Accounting journals, and have a good teaching record. About the university and school: Please visit: http://www.nbs.ntu.edu.sg/info.pdf. Application procedure: Interested individuals are encouraged to email their cover letter, *curriculum vitae*, personal particulars (http://www.ntu.edu.sg/hr/forms/PPF.doc) working papers, teaching record, and three letters of recommendation to: Professor Hun-Tong Tan, Chairman, Faculty Search Committee, Division of Accounting, Nanyang Business School, Nanyang Technological University S3-01B-51, Nanyang Avenue, Singapore 639798. Email: NBS_Search@ntu.edu.sg.

316 UNIVERSITY OF IDAHO, Department of Accounting seeks applicants for tenure-track faculty appointment as Assistant or Associate Professor of Accounting beginning August 2008 or January 2009. Teaching emphasis in Financial Accounting. Salary depends on qualifications and experience. Ph.D. in Accounting required by appointment date. Candidates should have demonstrated interest and/or record of teaching, research, and service accomplishments. Professional certification desired. Applicants should apply online at: http://www.hr.uidaho.edu (job seekers) or send applications including a complete resume, a graduate studies transcript, teaching evaluations, and letters from three professional references to: Dr. Teresa P. Gordon, Department of Accounting, University of Idaho, PO Box 443161, Moscow, ID 83844-3161; Phone: (208) 885-6453. Screening will begin immediately. AA/EOE.

317 NORTH DAKOTA STATE UNIVERSITY, College of Business (AACSB) invites applications for a tenure-track position in the Department of Accounting and Information Systems effective either August 15, 2008 or January 1, 2009. Rank is open depending on qualifications. Teaching responsibility can be in the financial or managerial accounting area. Minimum qualifications: Ph.D. or D.B.A. in accounting with research and teaching potential (ABD will be considered), good interpersonal skills, demonstrated effective written and verbal communication skills. Preferred qualifications: proven record of publications and teaching. Application procedure: review of applications begins March 31 and continues until the position is filled. Send application letter, *vita*, summaries of teaching performance and name, address, phone number of at least three references to: Dr. Herb Snyder, Search Committee Chair, PO Box 5075, NDSU, Fargo, ND 58105; or email: herbert.snyder@ndsu.edu. For more information concerning the position: http://www.ndsu.edu/ndsu/jobs/non_broadbanded/index.shtml. Additional information concerning the NDSU College of Business and the City of Fargo may be found at: http://www.ndsu.nodak.edu/business/ and http://www.ci.fargo.nd.us/VisitingFargo/. NDSU is an Equal Opportunity Institution.